

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 1, 2003, 08:50:33 ; Search time 3972 Seconds
(without alignments)
2770.566 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433

Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPKGEAKPRKRVRRPQR 269

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xih
-Q=/cgn2.1/USPTO.spool/US10049742/runat_01122003_073136_18605/app_query.fasta_1.455
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10049742@cgn2.1 3508 @runat_01122003_073136_18605 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGIOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rnd.*
- 36: em_htg_mam.*
- 37: em_htg_vit.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1433	100.0	1147	6	AX081438	AX081438 Sequence
2	1433	100.0	2343	6	AX713973	AX713973 Sequence
3	1433	100.0	2343	9	AK055945	AK055945 Homo sapi
4	1433	100.0	4274	9	BC050271	BC050271 Homo sapi
5	1397	97.5	1818	9	BC016941	BC016941 Homo sapi
6	1367	95.4	2402	4	AY027882	AY027882 Bos tauru
7	1367	95.4	3159	4	AY027881	AY027881 Bos tauru
8	1365	95.3	2392	10	AF351783	AF351783 Rattus no
9	1361	95.0	2692	4	AF308815	AF308815 Bos tauru
10	1357	94.7	2402	10	BC011146	BC011146 Mus muscu
11	1306	91.1	26371	2	AC098454	AC098454 Rattus no
12	1201	83.8	190201	2	AC129837	AC129837 Papio anu
13	1143	79.8	953	9	AF351784	AF351784 Homo sapi
14	1112	77.6	747	6	BD020639	BD020639 Novel Gen
15	1112	77.6	747	6	BD100577	BD100577 Novel Gen
16	1051	73.3	754	9	AF141342	AF141342 Homo sapi
17	1012.5	70.7	118052	9	AC023055	AC023055 Homo sapi
18	1004.5	70.1	247309	2	AC097837	AC097837 Rattus no
19	777.5	54.3	1253	14	AF268178	AF268178 Bovine vi
20	766	53.5	1258	14	AF268176	AF268176 Bovine vi
21	699	48.8	192658	2	AC131081	AC131081 Mus muscu
22	696.5	48.6	245462	2	AC097931	AC097931 Rattus no
23	693	48.4	927	14	BDU43603	BDU43603 Border dise
24	689.5	48.1	1233	14	AF268172	AF268172 Bovine vi
25	685	47.8	1004	14	AF268171	AF268171 Bovine vi
26	618.5	43.2	752	14	AF268177	AF268177 Bovine vi
27	606	42.3	13203	14	AB078952	AB078952 Bovine vi
28	601	41.9	13196	14	AB078951	AB078951 Bovine vi
29	586.5	40.9	873	14	BDU43602	BDU43602 Border dise
30	571	39.8	1229	14	AF268180	AF268180 Bovine vi
31	565	39.4	717	14	BVD1251N1	BVD1251N1
32	561	38.1	852	14	AF268179	AF268179 Bovine vi
33	549	38.3	711	14	BVD1251N2	BVD1251N2
34	534	37.3	12602	14	AF144617	AF144617 Pestiviru
35	506.5	35.3	211163	2	AC141508	AC141508 Rattus no
36	498	34.8	1145	14	BVD25053	BVD25053
37	497	34.7	12119	6	AX057315	AX057315 Sequence
38	497	34.7	12573	14	BVDCG	BVDCG
39	497	34.7	12578	14	BV133738	BV133738 Bovine vi
40	497	34.7	12734	6	AR179057	AR179057 Sequence
41	497	34.7	12734	14	AF268278	AF268278 Pestiviru
42	497	34.7	14078	6	AR123006	AR123006 Sequence
43	497	34.7	14078	6	AR214645	AR214645 Sequence
44	497	34.7	14578	6	AR094147	AR094147 Sequence
45	493.5	34.4	1114	14	AF268175	AF268175 Bovine vi

ALIGNMENTS

RESULT 1

AX081438
LOCUS AX081438 1147 bp DNA linear PAT 27-FEB-2001
DEFINITION Sequence 22 from Patent WO0109178.
ACCESSION AX081438
VERSION AX081438.1 GI:13170261
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue, H., Bandman, O., Tang, Y. T., Baughn, M. R., Azimzai, Y. and Lu, D. A.
TITLE Human chaperone proteins
JOURNAL Patent: WO 0109178-A 22 08-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..1147
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 1459372CB1"
BASE COUNT 287 a 278 c 314 g 268 t
ORIGIN
Alignment Scores:
Pred. No.: 1.36e-91 Length: 1147
Score: 1433.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-049-742-11 (1-269) x AX081438 (1-1147)
QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
DB 98 ATGGCTGGGGTTCCTGAGGATGAGCTAAACCCCTTCCATGTACTGGGGTTGAGGCCACA 157
QY 21 AlaSerAspValCbluLeuLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
DB 158 GCATCAGATGTTGAATGAAGAGGCGCTATAGACAGCTGCGAGTGTGCTTCATCTGCAC 217
QY 41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
DB 218 AAAAATCATATCCCGGCTGAGGAGGCTTCAAGGTTTTCGAGCGACTTGGGACATT 277
QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
DB 278 GTCAGCAATGCTGAAAAGCGAAAGGAGTATGAGATGAACGATGCGACAGAAATGAGCTG 337
QY 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
DB 338 AGCGGTCAGTAAATGAGTTTCTGTCCAAGCTGCAAGATGACCTCAAGGAGGCAATGAAT 397
QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluPro 120
DB 398 ACTATGATGTGAGCGGATGCCAAGAAAGCATAGAGGTTTGAATGGACCGGAACT 457
QY 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluClyAspPhe 140
DB 458 AAGAGTGCCAGATATCTGTGCTGAGTGTAATAGCTGCACTCTGCTGAGAAAGGACATT 517
QY 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
DB 518 TGGGCAGAGTCAAGCATGTGGGCTCAAGATCACTACTTTCATCTGATGGATGGAAG 577
QY 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
DB 578 GTGTATGATCATCACAGAGTGGGCTGGATGCCAGCGTGTAGGTATCTCCCCAGATACCCAC 637
QY 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
DB 638 AGAGTCCCTATCATCTCTATTGGTCTCGGATTCAGGACACAGAGCGCGCAGAGA 697

QY 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
DB 698 GCCACCCAGATGCCCTCCTGCTGATCTTTCAGGATTTCTTTCAGTTCGATCTTTCAAGTA 757
QY 221 ProProGlyGlnMetProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAla 240
DB 758 CCCCAGGCGAGATGCCCAATGGGAATCTTTCAGCTCCTCAGCTGCCCTGGAGCC 817
QY 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
DB 818 GCTCAGGCTCTAAGCCCAACAGCACAGTACCCAGGAGGAAGCAACCTTAAGCGCGG 877
QY 261 LysLysValArgArgProPheGlnArg 269
DB 878 AAGAAAGTGAGGAGGCCCTTCCAACGT 904
RESULT 2
AX713973 2343 bp DNA linear PAT 15-APR-2003
LOCUS AX713973
DEFINITION Sequence 657 from Patent EP1293569.
ACCESSION AX713973
VERSION AX713973.1 GI:29888901
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.
TITLE Full-length cDNAs
JOURNAL Patent: EP 1293569-A 657 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
FEATURES
Location/Qualifiers
1..2343
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 547 a 527 c 633 g 636 t
ORIGIN
Alignment Scores:
Pred. No.: 3.06e-91 Length: 2343
Score: 1433.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-049-742-11 (1-269) x AX713973 (1-2343)
QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
DB 542 ATGGCTGGGGTTCCTGAGGATGAGCTAAACCCCTTTCATGTACTTGGGGTTGAGGCCACA 601
QY 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
DB 602 GCATCAGATGTTGAATGAAGAGGCGCTATAGACAGCTGCGAGTGTTCATCTGCCTGAC 661
QY 41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
DB 662 AAAAATCATATCCCGGCTGAGGAGGCTTCAAGGTTTTCGAGCAGCAGCTTGGACATT 721
QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
DB 722 GTCAGCAATGCTGAAAAGCGAAAGGAGTATGAGATGAACGAAATGCGAGAAATGAGCTG 781
QY 81 SerArgSerValAsnGluPheLeuSerLysLysLeuGlnAspLeuLysGluAlaMetAsn 100

Db 782 AGCCGGTCAGTAATGATTTCTTCCCAAGCTGCAAGATGACCTCAAGGAGGCAATGAAT 841
Qy 101 ThrMetMetCysSerArgCysGlnGlyHisArgPheGluMetAspArgGluPro 120
Db 842 ACTATGATGTCAGCCGATGCAAGGAAAGCATAGGAGGTTTGAATGGACCGGAACCT 901
Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
Db 902 AAGAGTGCCAGATAGTGTGCTGAGTGTAATAGGCTGCATCTGCTGAGGAGGAGACTTT 961
Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 962 TGGGACAGATCAGCATGTTGGGCTCAAGATCACCTACTTTGACATGATGGATGGAAG 1021
Qy 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
Db 1022 GTGTATGACATCAGCATGCGCTGATGCCAGGCTGTAGTATCTCTCCACAGATACCCAC 1081
Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
Db 1082 AGAGTCCCTATCATCATCTCATTTGGTTCTCGGATTCAGGACCAAGGAGGCGGACAGA 1141
Qy 201 AlaThrProAspAlaProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
Db 1142 GCCACCCAGATGCCCTCTCTGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTA 1201
Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyVala 240
Db 1202 CCCCAGGGGAGATGCCCAATGGAACTTCTTGCAGCTCTCTAGCTGCCCTGGAGCC 1261
Qy 241 AlaAlaIleSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
Db 1262 GCTGCGCTCTTAAGCCCAAGCAGCACAGTACCCCAAGGAGGAGCAACCTAAGCGCGG 1321
Qy 261 LysLysValArgArgProPheGlnArg 269
Db 1322 AAGAAAGTGAGGAGGCGCTTCCACGT 1348

RESULT 3

AK055945 2343 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ1383 fis, clone NHNP2000206, highly similar
DEFINITION to Bovine viral diarrhea virus type 2 strain BVDV2-SD1630C
polyprotein gene.
AK055945
AK055945.1 GI:16550797
VERSION oligo capping; fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

Eukaryota: Metazoa; Chordata; Craniata: Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Makamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuko,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished

2 (bases 1 to 2343)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES

source

Location/Qualifiers

1..2343
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NHNP2000206"
/cell_type="normal neural progenitor cells (NHNP5958)"
/clone_lib="NHNP2"
/note="cloning vector: pME18SFL3-primary culture, normal
neural progenitor cells"
113..1351
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71050.1"
/db_xref="GI:16550798"
/translation="MGVMTGRLLGQWQVNFQSLQSGFYCGVGLTRFLKLLGALLLLA
LALFLGLQWRFVGLGDRGLWDLKATWFLSDSPALQRLTLRDRSPQWLVR
IVQWGLEUPVYKONINROGNAPVAGRGYQPEEVARLLTWAGVPELNPVHVLGV
EATASDELKXAYROLAVMYHPDKNHPRAEAFKVLRAMDIVSNAEKRYEKMKRM
ANELUSRVNEFLSKLQDLKEAMTMCSRCOGHRRFEMDREFAKRYCAEKNRLH
PAEEDFWAESMLGLKITFYFALMDGKVYDITWAGCQGVGSPDTHRVPHISFGR
IPGTRGRATPDAPPADLQDFLSRIFQVPPQMPNGNFFAAPPAPGAAAAAKPNST
VPKGEAKPKRRKKVRPPQR"

BASE COUNT 547 a 527 c 633 g 636 t

ORIGIN

Alignment Scores:

Pred. No.: 3,06e-91 Length: 2343
Score: 1433.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-049-742-11 (1-269) x AK055945 (1-2343)

Qy 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
Db 542 ATGGCTGGGGTTCCTTGAGGATGAGCTAAACCTTTCCATGTACTGGGGTTGAGGCCACA 601
Qy 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
Db 602 GCATCAGATGTTCAACTGAAGAAGCGCTATAGACAGCTGGCAGTGATGTTTCATCTGAC 661
Qy 41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
Db 662 AAAAAATCATCATCCCCGGGCTGAGGAGCGCTTCAAGGTTTTCGAGCAGCTTGGGACATT 721
Qy 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
Db 722 GTCCAGCATGCTGAAAGCGAAGAGATGATGAGATGAAACGAATGGCAGAGATGAGCTG 781
Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
Db 782 AGCCGGTCAGTAATAGAGTTTCTGTCCAAGCTGCAAGATGACCTCAAGGAGGCAATGAAT 841
Qy 101 ThrMetMetCysSerArgCysGlnGlyHisArgPheGluMetAspArgGluPro 120
Db 842 ACTATGATGTGTAGCCGATGCCAAGAAAGCATAGGAGGTTTGAATGGACCGGAACCT 901
Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
Db 902 AAGAGTGCCAGATCTGTGCTGAGTGTATAGGCTGCATCTCTGCTGAGGAGGAGACTTT 961
Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 962 TGGGACAGATCAAGCATGTTGGGCTTCAAGATCACCTACTTTGCACTGATGGATGGAAG 1021
Qy 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180

Db 1022 GTGATGACATCACAGAGTGGCTGGATGCCAGCGTGTAGTATCTCCCCAGATACCCAC 1081

Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyValArgGlnArg 200

Db 1082 AGAGTCCCTATCATCTCATTTGGTTCTCGGATTCAGGACACAGAGGGCGGCACAGA 1141

Qy 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220

Db 1142 GCCACCCAGATGCCCTCTGCTGATCTTCAGGATTTCTTGAGTCCGATCTTTCAAGTA 1201

Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240

Db 1202 CCCCAGGCGAGATGCCAAATGGGAACCTCTTTTCAGCTCTCCAGCTGCCCTGGAGCC 1261

Qy 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260

Db 1262 GCTGAGGCTCTTAGCCCAACAGACAGTACCCAGGAGGAAGCCAAACCTTAAGCGGG 1321

Qy 261 LysLysValArgArgProPheGlnArg 269

Db 1322 AAGAAAGTGAGAGGCCCTTCCACGT 1348

RESULT 4

BC050271

LOCUS

DEFINITION Homo sapiens, Similar to RIKEN cDNA 5730551F12 gene, clone

ACCESSION BC050271.1 GI:29792138

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 22 Row: 1 Column: 8

This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.

Location/Qualifiers

1. .4274

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:17391 IMAGE:3913714"

/tissue_type="Uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_71"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

93..2201

/codon_start=1

/product="Similar to RIKEN cDNA 5730551F12 gene"

CDS

/protein_id="AAH50271.1"

/db_xref="GI:29792138"

/translation="MAQKHPGSRGLYGAHSGGASLRTLGPVDEIPSPFSGLRDSAG

TPANGTRCTEHSQPKHTQNPAPHLWDSPGPGPPDAEDPDQSTSESEESG

VQBELSKENETGNQKDGNSFLSPSACNCOGTGGIPEGYSSEGGNGSSNFCCHCTSP

ALGEDLESEYDDESLKFPSPDFSVSSCKQKPPSRQRHRTFTKTRGGRRDPRSP

GHRHGRKESQADKRGKLGWABELCQIQAGFWMLIELLVIGVEYVTCGHLIYAC

ROLKSDLDLFRVMVGTDRGLGWAQVNFQFSQGYCGVGLFTFLKLLGALLLA

LALFLQFLQWRFLVGLDRGWRDKATWLFSLWDS PALQRLTLRLRSRQWLVR

IVQGWLETPVTKQNTNROGNAPVASGRYCOPEEEVARLLTMAGVPEDELNPVHLGV

EATASDVEUKKAYRQLAVNMHEDKNNHPRAEAFKVLRAANDIVSNAEKREYEMKRM

ANELUSRYNEFLSKLQDDLEKNNMCMSCOGKXRRFEMDREPKSARYCAECNKLH

PAEEDFWAESMLGKITFYALMDGKVYDITWAGCQKRGVISPTRHYVYHISGRS

ICGTGRQRATPDAPPADLQDFLSRIFQVPPQMPNGNFFAAPPAPGAAGAASKPNST

VFKGAKPKRRKKVRPPQR"

BASE COUNT 1065 a 1101 c 1143 g 965 t

ORIGIN

Alignment Scores:

Pred. No.: 6 05e-91 Length: 4274

Score: 1433.00 Matches: 269

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-10-049-742-11 (1-269) x BC050271 (1-4274)

Qy 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20

Db 1392 ATGGCTGGGGTTCTCGAGGATGAGCTAAACCTTTCCATGTTACTGGGGTTGAGGCCACA 1451

Qy 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40

Db 1452 GCATCAGATGTTGAACCTGAAGAAGGCCTATAGACAGCTGGCAGTGTGCTTCTATCCTGAC 1511

Qy 41 LysAsnHisHisProArgAlaGluAlaPheLysValLeuArgAlaThrAspIle 60

Db 1512 AAAAATCATCATCCCGGCTGAGGAGGCCTTCAAGGTTTTCGAGCAGCTTGGGACATT 1571

Qy 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80

Db 1572 GTCAGCAATGCTGAAAGCGAAGGAGTATGAGTGAACGAAATGCCAGAGATGAGCTG 1631

Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100

Db 1632 AGCCGTCAGTAAATGAGTTTCTGTCGAAGCTGCAAGATGACCTCAAGAGGCAATGAT 1691

Qy 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120

Db 1692 ACTATGATGTGTAGCCGATGCCAAGGAAAGCATAGGAGGTTTCAAATGACACCGGAACT 1751

Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLysLeuHisProAlaGluGluGlyAspPhe 140

Db 1752 AAGAGTGCCAGATCTGCTGAGTGTAAATAGGCTGCATCTCTGTCAGGAGGAGACTTT 1811

Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160

Db 1812 TGGGCAAGTCAAGCATGTTGGGCTTCAAGATCACCTACTTTGCACTGATGATGAAAG 1871

Qy 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180

Db 1872 GTGTATGACATCACAGATGGGCTGGATGCCAGCGTGTAGGTATCTCCCAATACCCAC 1931

Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200

Db 1932 AGAGTCCCTATCATCTCATTTGGTTCTCGGATTCAGGACACAGAGGGCGGCACAGA 1991

Qy 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220

Db 1992 GCCACCCAGATGCCCTCTGCTGATCTTCAGGATTTCTTGGTGGATCTTTCAAGTA 2051

Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240

Db 2052 CCCCCAGGAGATGCCAATGGGAACCTCTTTTCAGCTCCTCAGCCTGCCCTGGAGCC 2111
Qy 241 AlalalaSsrLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
Db 2112 GCTCAGCCTCTAAGCCCAACAGACAGTATACCCCAAGGAGAGCAAACTAAGCGCGG 2171
Qy 261 LysLysValArgArgProPheGlnArg 269
Db 2172 AAGAAAGTGAGGAGGCCCTTCCACGT 2198

RESULT 5
LOCUS BC016941 1818 bp mRNA linear PRI 09-NOV-2001
DEFINITION Homo sapiens, clone MGC:21452 IMAGE:3448446, mRNA, complete cds.
ACCESSION BC016941
VERSION BC016941.1 GI:16877383
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1818)
Strausberg,R.
Direct Submission
Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 20 Row: a Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES
source
1. .1818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:21452 IMAGE:3448446"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
401..1570
/codon_start=1
/product="Unknown (protein for MGC:21452)"
/protein_id="AAH16941.1"
/db_xref="GI:16877384"
/translations="MKRMENELSRVNEFLSKLQDDLKEAMNTMCSRQGHRRFE
MDREPKARCAENLRHFAEGDFWAEESMLGLKITYFALMDGVYDITETWAGQORV
GISPDTPHYHISFGSRIPGTRGRQATPDAPPADLQDFLSRIFOVPPGMPNGNFF
AAQAPAGAAASKPNSTVPKGAEPKRRKLAEKOECLARGLETGKIKDILHRLQ
AYLEHAEAEENEEDVLGDETEBEETKPIELPVKEEPEPEKTVDVAAEKYVKIISEI
PQTERVQKAEERNVPSVLESKYARAAFGISSVFKGLSSDNKPMVNDLKKERRAQ
RFLINVSKRSKDEDEKLKKRRERGIIVTSAGTGTEDTEAKRKRERFGIA"
532 a 373 c 494 g 419 t

BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 7,57e-89 Length: 1818
Score: 1397.00 Matches: 262
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 97.49% Indels: 0
DB: Gaps: 0
US-10-049-742-11 (1-269) x BC016941 (1-1818)
Qy 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
Db 188 ATGGCTGGGGTTCCTGAGGATGAGCTAAACCCCTTTCCATGTACTGGGGTTGAGGCCACA 247
Qy 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
Db 248 GCATCAGATGTTGAACCTGAAGAAGGCCTATACACAGCTGCAGTGATGCTTCATCCTGAC 307
Qy 41 LysAsnHisHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
Db 308 AAAAATCATCATCCCCGGCTGAGGAGGCCTTCAAGGTTTTCGAGCAGCTTGGGCATTT 367
Qy 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
Db 368 GTCCAGCAATGCCTGAAAGCGAAGGAGTATGAGATGAAACGAAATGCCAGAGATGAGCTG 427
Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
Db 428 ASCCGGTCAAGTAAATGAGTTTCTGCCAAGCTGCAAGATGACCTCAAGGAGGCAATGAAT 487
Qy 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluPro 120
Db 488 ACTATGATGTGTAGCGGATGCCAAGAAAGCATAGAGGTTTGAATGGACCGGAACCT 547
Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPhe 140
Db 548 AAAGAGTGCCAGATACTGTGCTGAGTGTATAGGCTGCATCTCTCTGAGGAGGAGACTTT 607
Qy 141 TrpAlaGluSerSerMetLeuGlyLysLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 608 TGGGCAGAGTCAAGCATGTGTGGGCTCAAGATCACCTACTTTGCACCTGATGATGGAAG 667
Qy 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyLysLeuSerProAspThrHis 180
Db 668 GTGTATGACATCACAGATGGGCTGGATGCCAGCGTGTAGGTATCTCCCCAGATACCCAC 727
Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
Db 728 AGAGTCCCCCTATCACATCTCATTTGGTTCTCGATTCCAGGCACCAGAGCGCGGACAGA 787
Qy 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
Db 788 GCCACCCCGAGATGCCCTCTCTGATCTTCAGGATTTCTTGAGTCGGATCTTCAAGTA 847
Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
Db 848 CCCCAGGCGCAGATCCCAATGGGAACCTTCTTTGCAAGCTCTCTCAGCCTGCCCCCTGAGCC 907
Qy 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
Db 908 GTTGAGGCTCTTAAGCCCCAACACAGCATCCCAAGGAGGAGAACCAACCTTAAGCGCGG 967
Qy 261 LysLysVal 263
Db 968 AAGAAACCTT 976
RESULT 6
AY027882
LOCUS Bos taurus J-domain protein Jiv mRNA, complete cds.
ACCESSION AY027882
VERSION AY027882.1 GI:15777194
KEYWORDS


```

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
1..187
188..2287
/function="interacts with nonstructural protein 2 of
bovine viral diarrhea virus"
/notes="contains putative zinc finger domain"
/codon_start=1
/product="J-domain protein Jiv"
/protein_id="AAK28650.1"
/db_xref="GI:15777193"
/translation="MAQKHGPGGLCGAHSGASRLTLGSPVDPDILSFSLGRDSAG
SAPNCTRLCTHSSPKYTOPNPAPAHSDPSHPGPGPGPLAEDPDQSEASSRESGV
DOELSRENETGYQDDGNSFLISPTCNCQGTGPIPEGPISEGRDSSSNFCHCTSP
ALGDEELEGYEDEEPLKFPDSVSPKXPPAPRRQHRVPAKEDTREGRRDTPS
PGRHLGRSQADKFLGKLGWAEELCOLQAGFWLIELLVGEYVETCGHLIYA
CQLKSGDLCLLRVWVWAGRLGWAQWQFQLSQGFCYAGLFTFLRVLGALLLL
ALALLGLDLGMRFLVGLSLGRLGWRDKATWIFSLASPTQWRLILLRESNPQHLV
RIVQWGLELPMVKORTNRQANAPVAGRYCQPEEVARLLTMAGVDEDENLPHVLG
VEATSDVELKAYQLAVMPDPKXNHPRAEEAFKVLRAAWDIVSNPERKEYEMR
MAENELSNVEFLSQLOAMNTMCMSCQKRRRFENDREPKSARYCAECNRLHPAE
EGDFWAEESMLGLKITYPALMDKGVYDITWAGCORVGISPDTHRPVHIHFGSRMPG
TSGRQATPDAPPADLODFLSRIFOVPPQWNGNFFAAPPQPGFCAATAASKPNSTVPK
GEAKPRKKVRRPQR"
2288..3159
/notes="use of alternative polyadenylation site results in
3' UTR longer than in the sequence presented in GenBank
Accession Number AY027882"
BASE COUNT 755 a 759 c 895 g 750 t
ORIGIN

Alignment Scores:
Pred. No.: 1
Score: 1.77e-86 Length: 3159
Percent Similarity: 1367.00 Matches: 258
Best Local Similarity: 96.65% Conservative: 2
Query Match: 95.91% Mismatches: 5
DB: 95.39% Indels: 4
Gaps: 1

US-10-049-742-11 (1-269) x AY027881 (1-3159)

QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
DB 1490 ATGGCTGGGGTTCCTGAGGATGAGCTAAACCTTTTACAGTGTGTGGGGTTGAAGCCACA 1549

QY 21 AlaSerAspValGluLeuLysValAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
DB 1550 GCATCAGATGTTGAACTAAGAAAGCCCTATAGGAGCTGGCAGTGCATGTTTCATCTGAC 1609

QY 41 LysAsnHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTyrAspIle 60
DB 1610 AAAAATCATCTCTGCTGAGAGACCTTTCAAGTTTTCGGGGCAGCTTGGGACATT 1669

QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
DB 1670 GTCAGCAACCTGAAAGAGCGAAGGATATGAGATGAACAGATGGCAGAAATGAGCTG 1729

QY 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
DB 1730 AGCCGGTCAGTGAATGATGTTTCTGTCAAGCTGCAA-----GAAGCAATGAAT 1777

QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120
DB 1778 ACGATGATGTGACCCGATGCCAGGAAAGCATAGGAGGTTTGAATGGACCGGAACCT 1837

QY 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
DB 1838 AAGAGTCCAGATACTGTGCTAGTGAATAGGCTGCATCTCTGCTGAGGAAGTGACTTT 1897

QY 141 TrpAlaGluSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
DB 1898 TGGCAGAGTCAAGCATGTTGGGCTCAAAATACCTACTTTGGCTGATGGATGGAAAG 1957

```

```

QY 161 ValTyrAspIleThrGluTyrProAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
DB 1958 GTGTATGATATCACAGATGGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCAC 2017

QY 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
DB 2018 AGAGTCCCTTATCACTCTCATTTGGTTACGATGCCAGCACCATGTGGCGGCAGAGA 2077

QY 201 AlaThrProAspAlaProProAlaAspPheLeuGlnAspPheLeuSerArgIlePheGlnVal 220
DB 2078 GCTACTCCAGATGCCCTCTCTGCTGACCTTCAGATTTCTTGACCGGATCTTTCAAGTA 2137

QY 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
DB 2138 CCCCAGCGCCAGATGTCACACGCGAACTCTTGTGACCTCTCAGCGCCGCTGGGGCC 2157

QY 241 AlaAlaLysSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
DB 2198 ACTGCACGCTCCAGCCCAACAGCACAGTACCAAGGAGAGCAACCAAGCGCGCG 2257

QY 261 LysLysValArgArgProPheGlnArg 269
DB 2258 AAGAAAGTGAGGAGGCCCTTCCACAGT 2284

```

```

RESULT 8
AF351783 LOCUS 2392 bp mRNA linear ROD 24-MAY-2001
DEFINITION Rattus norvegicus dopamine receptor interacting protein (Drip78)
mRNA, complete cds.
ACCESSION AF351783
VERSION AF351783.1 GI:14194054
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2392)
AUTHORS Bermak, J.C., Li, M., Bullock, C. and Zhou, Q.-Y.
TITLE Regulation of transport of the dopamine D1 receptor by a new
membrane-associated ER protein
JOURNAL Nat. Cell Biol. 3 (5), 492-498 (2001)
MEDLINE 21231375
PUBMED 11331877
REFERENCE 2 (bases 1 to 2392)
AUTHORS Bermak, J.C., Li, M., Bullock, C.M. and Zhou, Q.-Y.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2001) Pharmacology, University of California,
1982 Jamboree Blvd., Irvine, CA 92697, USA
FEATURES
source
1..2392
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/notes="extrachromosomal"
1..2392
/gene="Drip78"
161..2266
/gene="Drip78"
/notes="Drip78; ER-membrane protein; zinc finger motif"
/codon_start=1
/product="dopamine receptor interacting protein"
/protein_id="AAK56240.1"
/db_xref="GI:14194055"
/translation="MAQKHGPGGLCGVHSGGSLITSGSSVDPEILSFSLGRDSKE
TAPNCTRLCTHSDPKCTQPNPAHSDPSHPGPGPGPLAEDPDQSEASSRESGV
DOELSRENETGYQDDGNSFLISPTCNCQGTGPIPEGPISEGRDSSSNFCHCTSP
ALGDEELEGYEDEEPLKFPDSVSPKXPPAPRRQHRVPAKEDTREGRRDTPS
PGRHLGRSQADKFLGKLGWAEELCOLQAGFWLIELLVGEYVETCGHLIYA
CQLKSGDLCLLRVWVWAGRLGWAQWQFQLSQGFCYAGLFTFLRVLGALLLL
ALALLGLDLGMRFLVGLSLGRLGWRDKATWIFSLASPTQWRLILLRESNPQHLV
RIVQWGLELPMVKORTNRQANAPVAGRYCQPEEVARLLTMAGVDEDENLPHVLG
VEATSDVELKAYQLAVMPDPKXNHPRAEEAFKVLRAAWDIVSNPERKEYEMR
MAENELSNVEFLSQLOAMNTMCMSCQKRRRFENDREPKSARYCAECNRLHPAE
EGDFWAEESMLGLKITYPALMDKGVYDITWAGCORVGISPDTHRPVHIHFGSRMPG
TSGRQATPDAPPADLODFLSRIFOVPPQWNGNFFAAPPQPGFCAATAASKPNSTVPK
GEAKPRKKVRRPQR"

```

IOWGWELEFPWKQRTQROGTANVPSGRYCPQEEVEVARLLTMAGVPEDELNPFHVLGV		565 a		582 c	723 g	522 t							
ATASDIENKAYRQIAVVMVHDKNHHPRAEEAFKVLRAADIVSNPERRKEYEMKRWMA													
ENELSRVNEFLSKLQDDKEMRMTMCSRCQGHRRFEMDRPEKRSARYCAECNRLHP													
ABEGDFWASSMLGKITFYALMDGKVYDITWAGCORVGI SPDTHRVYHISFGSRV													
PGTSGRQATPESPPADLQDFLSRIFQVPPGPMNSNGNFFAAPHFGPGTSTSRPNSSV													
PKGEAKPRKRRPFRQ"													
BASE COUNT		565 a		582 c		723 g		522 t					
ORIGIN													
Alignment Scores:		1.79e-86		Length:		2392							
Pred. No.:		1365.00		Matches:		252							
Score:		96.65%		Conservative:		8							
Percent Similarity:		93.68%		Mismatches:		9							
Best Local Similarity:		95.25%		Indels:		0							
Query Match:		10		Gaps:		0							
DB:													
US-10-049-742-11 (1-269) x AF351783 (1-2392)													
Qy	1	MetAlaGlyValProGluAspGluLeuAenProPheHisValLeuGlyValGluAlaThr	20										
Db	1457	ATGGCAGGGGTTCTGAAGATGAACCTAAACCCCTTTTCATGCTGGGGTTGAAGCTACA	1516										
Qy	21	AlaSerAspValGluLeuLysLysAlaTyArgGlnLeuAlaValMetValHisProAsp	40										
Db	1517	GCATCCGATATTGAACATAAGAGGCGCTATAGCAGCTAGCGGTGATGCTCATCTCGAT	1576										
Qy	41	LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle	60										
Db	1577	AAAAACCATCCCGCGGTGAGGAGGCGCTTTAAAGTTTTCGGGCGAGCTGGGACATC	1636										
Qy	61	ValSerAsnAlaGluLysArgLysGluTyArgGlnLeuMetLysArgMetAlaGluAsnGluLeu	80										
Db	1637	GTCCAGCAATCCAGAGAGCGGAGGAATATGAGATGAACCGATGCGAGAGATGAGCTC	1696										
Qy	81	SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn	100										
Db	1697	AGCCGTCGGTGAATGAGTTCTGTCAGAGCTACAGATGACCTCAGAGAGCAATGAT	1756										
Qy	101	ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro	120										
Db	1757	ACGATGATGTGCAGCGCGTGCCAGGGAAGCATAGGAGGTTTGAATGACCGGAACCC	1816										
Qy	121	LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPhe	140										
Db	1817	AAGAGTGCCAGATACCTGTCGAGTGTAAACAGTTGCTGCTGAGGAGAGAGACTTC	1876										
Qy	141	TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys	160										
Db	1877	TGGCAGAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTGCACCTGATGATGGAG	1936										
Qy	161	ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis	180										
Db	1937	GTGTATGATCATCACAGTGGGTGATGCCAGCGTGTGGGATCTCCCCAGATACCCAC	1996										
Qy	181	ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg	200										
Db	1997	CGAGTTTCTTACCACATCTCATTTGGTTCTCGGTCACAGGCACCGCCGCAAAAGG	2056										
Qy	201	AlaThrProAspAlaProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal	220										
Db	2057	GCCACTCCAGAGTCCCTCTGCTGACCTGCAGGATTTCTTGAGCCGGATCTTTCAAGTA	2116										
Qy	221	ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla	240										
Db	2117	CCTCCAGGCGCGATGTCGAATGGGAACCTCTTTGCCGCACCTCACCTCGCCCTGGGACT	2176										
Qy	241	AlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg	260										
Db	2177	ACTTCAACCTCCAGGCCCAACAGTTCTAGTACCCAGGAGAGCCAAACCTAAACGGCGG	2236										
Qy	261	LysLysValArgArgProPheGlnArg	269										

Db

2237

AAGAAGTGAAGCGGCCCTTCCAAGA

2263

RESULT 9

AF308815

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

FEATURES

source

CDS

AF308815

Bos taurus DnaJl protein mRNA, partial cds.

AF308815

AF308815.2

GI:15843560

Bos taurus (cow)

Bos taurus

1 (bases 1 to 2692)

Neill,J.D. and Ridpath,J.F.

Recombination with a cellular mRNA encoding a novel DnaJ protein results in biotype conversion in genotype 2 bovine viral diarrhoea viruses

Virus Res. 79 (1-2), 59-69 (2001)

21436421

11551646

2 (bases 1 to 2692)

Neill,J.D. and Ridpath,J.F.

Direct Submission

Submitted (27-SEP-2000)

VPDLRU, National Animal Disease Center, 2300 Dayton Ave., Ames, IA 50010, USA

3 (bases 1 to 2692)

Neill,J.D. and Ridpath,J.F.

Direct Submission

Submitted (02-OCT-2001)

VPDLRU, National Animal Disease Center, 2300 Dayton Ave., Ames, IA 50010, USA

Sequence update by submitter

On Oct 2, 2001 this sequence version replaced gi:12584131.

Location/Qualifiers

1..2692

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

<1..1982

/note="bDnaJ1"

/codon_start=3

/product="DnaJl protein"

/protein_id="AAG59810.2"

/db_xref="GI:15843561"

/translation="SAGSAPNGTRCLTEHSSPKYTOPNPAHMSDPHGPPRGGPPLA
AEEPDQSPASSEGVDDLEARENETGYDDGNSSFI SIPTCNCQGTGFIGEPGYS
EGKDSNNFCHCTSPALGEDEBELLSFPVMFHVCPSEKKAQPRQRPR
VPKEDTREGRRDPRSPGRHLGRKRSQADKRGGLGMAELCOLQAGQFWLIEL
LVLEGNFRELVCALLALALLLGLCQLGWRFLRIEMTGGWRDKATWIFSWLASPT
AGLFRFLVGLALLLALALLLGLCQLGWRFLRIEMTGGWRDKATWIFSWLASPT
WQRLIILRESPPQQLVRIVQWGLELPWKQRTNRQANAPVAGGRYCPQEEVRLA
LTMAGVPEDELNPFHVLGVEATSDVELKAYROLAVMHPDKNHPRAEAEAFKVLRA
ANDVSNPERRKEYEMKMAENELSRVNEFLSKLOEAMNTMCSRCQGHRSFEMDR
EPKSAKYCACNRLHPAEEGDFWASSMLGKITFYALMDCKVYDITWAGCORVGIS
PDTHRVYHISFGSRVTSRQRATPDAPPDLQDFLSRI FQVPPGQMSNGNFFAAP
QPGGATAASKPNSTVPKGEAKPRKRRPFRQ"

BASE COUNT

670 a

635 c

756 g

631 t

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

3,89e-86

1361.00

96.28%

95.54%

94.98%

4

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

2692

257

2

6

94

1

US-10-049-742-11 (1-269) x AF308815 (1-2692)

Qy

1

MetaAlaGlyValproGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr

20

Db

1185

ATGGCTGGGGTCTTCAGATAGAGCTAAACCTTTTTCACGTGTGGGGTTGAAGGCACA

1244

Tue Dec 2 08:21:30 2003

QY	21	AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp	40
DB	1245	GCATCAGATTCTGAACCTGAAGAAGGCGCTTAGGCAGCTGGCAGTGATGGTTTCATCCTGAC	1304
QY	41	LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle	60
DB	1305	AAAAATCATCATCTCTGTCGGAGGAAGCTTTCAAGGTTTTCGGCGCAGCTTCGGACATT	1364
QY	61	ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu	80
DB	1365	GTCAAGACCCCTGAAGACCGAAGGAATATGAGATGAAACGANTGGCAGAAAAATGAGCTG	1424
QY	81	SerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsn	100
DB	1425	AGCCGGTCAGTGAATGAGTTCTGTCCAAGCTGCAA-----GAAGCAATGAAT	1472
QY	101	ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro	120
DB	1473	ACGATGATGTCAGCCGATCCACGGGAAAGCATTAGGAGCTTTGAAATGGACCGGAACCT	1532
QY	121	LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluArgAspPhe	140
DB	1533	AAAGAGTGCACAGTACTGTGCTGAGTGAATAGAGCTGCATCTCTGCTGAGGAAGTGACTTT	1592
QY	141	TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys	160
DB	1593	TGGCGAGAGTCAAGCATGTTGGCGCTCAAAATCACCTACTTTGCGCTGATGGATGGAAG	1652
QY	161	ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis	180
DB	1653	GTGTATGATATCACAGATGGCGTGGATGCCAGCGTGTGGGATCTCCCAGATATCCAC	1712
QY	181	ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg	200
DB	1713	AGAGTCCCTTATCACATCTCATTTGGTTTCAGGATGCCAGGCACCACTGGCGGCAGAGA	1772
QY	201	AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal	220
DB	1773	GCTACTCCAGATGCCCTCTCTGTGACCTTCAGGATTTCTGTAGCCGGATCTTTCAAGTA	1832
QY	221	ProProGlyGlnMetProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAla	240
DB	1833	CCCCCAGCCAGATGTCAAACGGGAACCTTTTTCAGCTCTCTCAGCCCGCCCTCGGGCC	1892
QY	241	AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg	260
DB	1893	ACTGCGAGCTCTCAAGCCCAACAGACAGTATCCCAAGGAGAACCCAAACGAGCGCGG	1952
QY	261	LysLysValArgArgProPheGlnArg	269
DB	1953	AAAGAAAGTGAGGAGGCGCTTTTCCAACGT	1979

RESULT 10	BC011146	2402 bp	linear	ROD 20-SEP-2002
LOCUS	BC011146			
DEFINITION	Mus musculus, RIKEN cDNA 573055F12 gene, clone MGC:19282			
ACCESSION	BC011146			
VERSION	BC011146.1	GI:15029845		
KEYWORDS	IMAGE:4016209, mRNA, complete cds..			
SOURCE	Mus musculus			
ORGANISM	Mus musculus (house mouse)			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
TITLE	1 (bases 1 to 2402)			
JOURNAL	Strausberg, R.			
	Direct Submission			
	Submitted (25-JUL-2001) National Institutes of Health, Mammalian			
	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
	USA			
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			

COMMIT

Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P. H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 24 Row: a Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction.

FEATURES	SOURCE
-----------------	---------------

```

FEATURES             Location/Qualifiers
     1..2402
         /organism="Mus musculus"
         /mol_type="mRNA"
         /strain="CZECH II"
         /db_xref="taxon:10090"
         /clone="MGC:19282 IMAGE:4016209"
         /issue_type="Mammary tumor metastatized to lung,
MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR
enhancer."
         /clone_lib="NCI CGAP_Lu30"
         /lab_host="DH10B"
         /note="Vector: pCMV-SPORT6"
     148..2259
         /codon_start=1
         /product="RIKEN cDNA 5730551F12 gene"
         /protein_id="AAH1146.1"
         /db_xref="GI:15029846"
         /db_xref="LocusID:74330"
         /translation="MAQHPGRRLCGAHRSGGTSLTSGSSVDPEILSFSGLRDSAE
TAPNCTRLCKEHSQPKYTPPNPAHWSDFHGPPRGSGVYDESETQSEESGV
DELSRENTGQEGCSFSFLGIPACNQCQSGFQVPEGYISEGSGSSSLCHCHCTSP
ALGEDELEEEYDDEEPLFPDSFDSVSGKKPLPRQKRRFPKEDVDSGRREPKA
PPHRRLARKSTDKRGLGLMGVEELQGLQAQFWKLIPELLVLGVEYVTCGLIYA
CKRLAGSDUDLFRVYVGMARLGGWARMQFLSQSFFCVVGLGILIRLUVVGAFLLL
ALALFGCLGFCVSGLGNRLGWRDKTAMFLSWLGSPLAHJCLTLTKDPSFWQQLV
RLIQWQDTELLKAVTQKKGNAVASGRYCOPEEVTRLLTMAGVPEDELANPFLVGL
VRATASDTELLKAVQLA/VMHDPKXNHHPRAEAFKILPAANDIVSNPERRKEYEMKR
MAENLSRSVNEFLSKLQDLLKEAMNTMCSQCQKRRFEMDRFPKRSARYCAECNRL
HPAEGDPFAEESMGLKITYFALMDKGVYDITEWAGCORVGISDTHTRVRYPTISFGS
RVFGTSGRQATYFESGPPVLDQFLSLRIFOVPEPNSNGNFFAAHPHGPCTTSTRPNS
SVYKGEAKPRKKRYRRFPQR"
     618 a      575 c      698 g      511 t
BASE COUNT

```

BASE CO
ORIGIN

Alignment Scores:

Pred. No.:	6.51e-86	Length:	2402
Score:	1357.00	Matches:	250
Percent Similarity:	95.91%	Conservative:	8
Best Local Similarity:	92.94%	Mismatches:	11
Query Match:	94.70%	Indels:	0
DB:	10	Gaps:	0

US-10-049-742-11 (1-269) x BC011146 (1-2402)

Qy	1	21	1510
MetAlaGlyValProGluLeuAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr	20	40	1569
1450	ATGCCTGGGGTTCCTGAAGATGAACATAAACCCCTTTTCATGTCTGGGGGTGAAGCTACA		1509
Db			
Qy	21	1510	
AlaSerAspValGluLeuLysAlaTyrArgGlnLeuAlaValMetValHisProAsp	40	1569	
1450	GCATCCGACACTGAACATAAGAAAGGGCTATAGCAGCATAGCAGTAATGTGTCATCCCTGAT		1509
Db			

Qy 41 LysAsnHisProArgAlaGluAlaPheLysValIleuArgAlaAlaTrpAspIle 60
 Db 1570 AAAAAATCACATCCCGGCTGAGGAGGCGCTTCAAAATTTTGGCGCAGCTTGGGACATT 1629
 Qy 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
 Db 1630 GTGAGCAACCCAGAGGCGGAGAGGAATATGAGATGAACGGATGCACAGATGAGCTC 1689
 Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
 Db 1690 AGCCGGTCAGTGAATGAGTTCTGTGGAACACTACAAGATGACCTCAAGGAGGCAATGAAC 1749
 Qy 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluPro 120
 Db 1750 ACATGATGTGCACGAGATCCCAAGGAAGACATAGGAGGTTTGAATGAGACCGGGAACCC 1809
 Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
 Db 1810 AAGAGTGCCAGATCTGTGCTGAGTGAACAGGCTGCTCTCTGCTGAGGAGAGACTTT 1869
 Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
 Db 1870 TGGCAGAGTCGAGCATGCTGGCCCTCAAGATCACATTACTTTCGGCTGATGGATGGAAG 1929
 Qy 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
 Db 1930 GTGTATGACATTACAGAGTGGCTGGATGCCAGGTGTGGGTATCTCTCCACAGATACTCAC 1989
 Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
 Db 1990 AGAGTTCCTTACCACATCTCATTTGGTTCCTCGGATACCCGGCACCTGTCGGCGAGG 2049
 Qy 201 AlaThrProAspAlaProProAlaAspPheLeuGlnAspPheLeuSerArgIlePheGlnVal 220
 Db 2050 GCCACTCCAGAGTCGCTCTCTGTGACCTGCAGGATTTCTTGAGCGGATCTTTCAGTA 2109
 Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
 Db 2110 CTTCCGGGGCGGATGTCCATGGGAATCTTTTGGCGCACCTCACCTGCGCTGGGACC 2169
 Qy 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
 Db 2170 ACTTCGACCTCTAGCGCCCAACAGTTTCAGTACCCCAAGGAGAGCAACCTAAACGGCGG 2229
 Qy 261 LysLysValArgArgProPheGlnArg 269
 Db 2230 AAGAAAGTGAGGCGGCTTTCACGA 2256
 RESULT 11
 AC098454/c 263371 bp DNA linear HTG 10-OCT-2002
 LOCUS Rattus norvegicus clone CH230-2H10, WORKING DRAFT SEQUENCE, 3
 DEFINITION unordered pieces.
 ACCESSION AC098454
 VERSION AC098454.4 GI:23664557
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 263371)
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshman,F.R., Allen,C., Alstrooms,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O. Denn,A.L., Ding,Y., Dinh,H.H., Douthwalte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Honsi,P., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,B., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,C., Lieu,C., Liu,J., Liu,W., Louisseged,H., Lozado,R.X., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,B., Moser,N., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,L., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sotnik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Woodson,C., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.
 Direct Submission
 TITLE JOURNAL
 REFERENCE 2 (bases 1 to 263371)
 AUTHORS Worley,K.C.
 TITLE JOURNAL
 REFERENCE 3 (bases 1 to 263371)
 AUTHORS Worley,K.C.
 TITLE JOURNAL
 COMMENT
 Submitted (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 Direct Submission
 Submitted (10-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Oct 10, 2002 this sequence version replaced gi:21953452.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: TUD
 Center clone name: CH230-2H10
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 243674 bases at least Q40
 Consensus quality: 246830 bases at least Q30
 Consensus quality: 248548 bases at least Q20
 Estimated insert size: 247731; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length

```

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 258056: contig of 258056 bp in length
* 258057 258156: gap of unknown length
* 258157 261384: contig of 328 bp in length
* 261385 261484: gap of unknown length
* 261485 263371: contig of 1887 bp in length.
* 263371 Location/Qualifiers
* 1..263371
  /organism="Rattus norvegicus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10116"
  /clone="CH230-2H10"
  misc_feature 1..1125
    /note="wgs contig"
  misc_feature 2121..4146
    /note="wgs contig"
  misc_feature 5480..6935
    /note="wgs contig"
  misc_feature 256941..258056
    /note="wgs contig"
  misc_feature 258157..259735
    /note="wgs contig"
  misc_feature 259889..261384
    /note="wgs contig"
  misc_feature 75390 a 46168 c 46816 g 81072 t 13925 others
  ORIGIN
Alignment Scores:
Pred. No.: 4,9e-80 Length: 263371
Score: 1306.00 Matches: 242
Percent Similarity: 93.68% Conservative: 10
Best Local Similarity: 89.96% Mismatches: 17
Query Match: 91.14% Indels: 0
DB: 2 Gaps: 0
US-10-049-742-11 (1-269) x AC098454 (1-263371)
QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
DB 143252 ATGGCTGGGGTCTCTGAGATGAACCTTTTTCATGCTGCTGGGGCTTGAAGCTACA 143193
QY 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
DB 143192 GCATCCGATATTGAACCTAAAGAAAGCCTATAGCAGCTAGCGGTGATGCTCCATCTGAT 143133
QY 41 LysAsnHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
DB 143132 AAAAATCACCATCCCGGCTGAGGAGGCTTTTAAAGTTTTCGGCGACGTTGGGACATT 143073
QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
DB 143072 GTCCAGCAATCCAGAGAGCGGAGGAATATGAGATGAACGATGCGACAGATGAGCTC 143013
QY 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
DB 143012 AGCCGCTCGGTGAATGAGTTTCTTCGAAGCTACAAGATGATTTCAGGAGGCAATGAT 142953
QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120
DB 142952 ACCATGATGTCAGCGGTGCCAAGGGAAGCATAGAGGTTTGAATGGACGGGAACCC 142893
QY 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
DB 142892 AAGAGTGCCAGATACCTGTGCTAGTGAACAGTTTTCATCTCTGAGGAGGAGACTTC 142833

```

```

QY 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
DB 142832 TGGGCAGAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTTGCACCTGATGGGAAG 142773
QY 161 ValTyrAspIleThrGluTyrAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
DB 142772 GTGTATGACATCAGAGAGTGTCTGGATGTCAGGCTGTGGGCACTCTCCCGACACCCAC 142713
QY 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
DB 142712 CAAGTTCCTTACCACACCTCATTTGGTTCTCGTTACCGGCACCGCGGGGCAACA 142653
QY 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
DB 142652 CCTCAGGCGCGATGTCATATGGGAACCTTTTCCCGCACCTCACCTTGGCGGACT 142533
QY 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
DB 142532 ACTTCAACCTCCAGGCCCAACAGTTTCAGTATCCAGGAGGAGGAGCAACCTAAAGCTGG 142473
QY 261 LysLysValArgArgProPheGlnArg 269
DB 142472 AAGAAAGTGAGTGGCGCTTCCAAAGCA 142446
RESULT 12
AC129837
LOCUS AC129837 190201 bp DNA linear HTG 11-JUN-2003
DEFINITION Papio anubis clone rp41-88j2, WORKING DRAFT SEQUENCE.
ACCESSION AC129837
VERSION AC129837.16 GI:30725963
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
REFERENCE 1 (bases 1 to 190201)
AUTHORS Prescott, A., Shaikh, T. and Roe, B.A.
TITLE Papio anubis BAC Clone rp41-88j2
JOURNAL Unpublished
2 (bases 1 to 190201)
REFERENCE Prescott, A. and Roe, B.A.
AUTHORS Direct Submission
TITLE Submitted (03-AUG-2002) Department Of Chemistry And Biochemistry,
JOURNAL The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 190201)
REFERENCE Prescott, A., Shaikh, T. and Roe, B.A.
AUTHORS Direct Submission
TITLE Submitted (11-JUN-2003) Department Of Chemistry And Biochemistry,
JOURNAL The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On May 15, 2003 this sequence version replaced gi:30698658.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 190201: contig of 190201 bp in length.
* Location/Qualifiers

```

FEATURES

```
source 1. .190201
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="xp41-88j2"
/clone_lib="pPCI - 41 Male (Olive) Baboon BAC Library"
BASE COUNT 46020 a 50800 c 49199 g 44180 t 2 Others
ORIGIN

Alignment Scores:
Pred. No.: 7.47e-73 Length: 190201
Score: 1201.00 Matches: 231
Percent Similarity: 90.71% Conservative: 13
Best Local Similarity: 85.87% Mismatches: 25
Query Match: 83.81% Indels: 1
DB: 2 Gaps: 0

US-10-049-742-11 (1-269) x AC129837 (1-190201)
Qy 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
Db 117839 TTGGCTGGGTTCCTGAGGATGAGTTAAACCTTTCATGTACTAGGCTTTGAGGCCACA 117898
Qy 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
Db 117899 GCATCAGACGCTGAACCTGAAGAAGGCTATAGACAGCTGCAGTGATGGTTTCATCCGCAC 117958
Qy 41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTyrAspIle 60
Db 117959 AAACTCATCCCCCGGCTGAGGAGCCTTCAAGTTTTTGGCGCAGTTTGGGACATT 118018
Qy 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
Db 118019 GTCAGCATGCTGAAGACGACAGGAATATGAGATGAAGAACGGCAGAGATCAGCTG 118078
Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsn 100
Db 118079 AGCCGGTCAGTAATAGATTCTCTCCAAAGTCGAAGATCACCTCAAGGAGGAGTGAAT 118138
Qy 101 ThrMetMetCysSerArgCysGlnGlyHisArgArgPheGluMetAspArgGluPro 120
Db 118139 ACTCTGATGTGGCGGATGCCAAGGAAGCATAGGAGTTTGAAGTGGACCAACCACT 118198
Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
Db 118199 AAGAGTCCAGATACTATGCTGCGTGTGATAGGCTGCGTCCCGCTGAGGAAGAGACTTT 118258
Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 118259 TGGCAGAGTCAGCTTGTGGGCTCAAGATCACCTTACTAGCACTGATGGATGGAAG 118318
Qy 161 ValTyrAspIleThrGluTyrAlaGlyCysGlnArgValGlyLysSerProAspThrHis 180
Db 118319 GTGTATGACACACACAGTGGGCTGGAGCGCAGCGCTCTAG-ATATCTCCCCAGACCCAC 118377
Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
Db 118378 AGAGTCTCCTACATCTCTCATTTGGTTCGGATTCTCAGGACCCAGGCGGCGCAGAGA 118437
Qy 201 AlaThrProAspAlaProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
Db 118438 GCCACCCGGAGCGCCCTCTCTGTGATCTTCAGGATTTCTTGAGTCAGATCTTCAAGTA 118497
Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
Db 118498 CACCCAGGCGAGCGCCAGTGGCAACTTCTTTCAGCTCTCTCAGCTCTCTCAGCGGCC 118557
Qy 241 AlaAlaLaserLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
Db 118558 ACTGACGCTTAAGCCCAACCGCACAGTATCCCAAGGAGAGAACGAAACCGAAGCGCGG 118617
Qy 261 LysLysValArgArgProPheGlnArg 269
```

Db 118618 ACGAAAGTAGGAGGCCCTTCCAAAGT 118644

RESULT 13
AF351784 953 bp mRNA linear PRI 24-MAY-2001
LOCUS Homo sapiens dopamine receptor interacting protein mRNA, partial cds.

ACCESSION AF351784
VERSION AF351784.1 GI:14194056

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 953)
AUTHORS Bermak,J.C., Li,M., Bullock,C. and Zhou,Q.-Y.

TITLE Regulation of transport of the dopamine D1 receptor by a new membrane-associated ER protein

JOURNAL Nat. Cell Biol. 3 (5), 492-498 (2001)

MEDLINE 21231375
PubMed 11331877

REFERENCE 2 (bases 1 to 953)
AUTHORS Bermak,J.C., Li,M., Bullock,C.M. and Zhou,Q.-Y.

TITLE Direct Submission

JOURNAL Submitted (21-FEB-2001) Pharmacology, University of California, 19182 Jamboree Blvd., Irvine, CA 92697, USA

FEATURES
source
1. .953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
c1. .644

CDS
/note="ORF178; ER-membrane protein; zinc finger motif"
/codon_start=3
/product="dopamine receptor interacting protein"
/protein_id="AAK56241.1"
/db_xref="GI:14194057"

translation="AWDIVSNAEKRYEMKMAENELSRVNEFLSKLQDLKEAMN
TMMSCCGKRRREMDEPKSARYCAECNLIHPAEGDFWAEISMLGIKITYPALMD
TKYDITITWAGCQNRVIGISPDTHRVPHISFGSRIPGTRGRQATDPADPADQLFLSR
IFQVPPGQMPNGVFFAAPPAPGAAAKSPNSTVPKGEAKPKRRKKRRPFR"

BASE COUNT 240 a 236 c 246 g 231 t

ORIGIN

Alignment Scores:
Pred. No.: 2.11e-71 Length: 953
Score: 1143.00 Matches: 213
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.78% Indels: 0
DB: 9 Gaps: 0

US-10-049-742-11 (1-269) x AF351784 (1-953)

Qy 57 AlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAla 76

Db 3 GTTGGGACATTGTGAGCAATGCTGAAAGCCGAAAGGAGTATGAGATGAACGAAATGGCA 62

Qy 77 GluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLys 96

Db 63 GAGATAGCTGAGCGGTCAGTAAATGAGTTCTGTCACAGCTGCAAGATGACCTCAAG 122

Qy 97 GluAlaMetAsnThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMet 116

Db 123 GAGGCAATGAATACTATGATGTGATAGCCGATCCCAAGGAACATAGGAGGTTTGAATG 182

Qy 117 AspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGlu 136

Db 183 GACCGGAACCTTAAGAGTGCCAGATCTGTGTAGTGTATAGGCTGTATAGGCTGCTGCTGAG 242

Qy 137 GluGlyAspPheThrAlaGluSerMetLeuGlyLeuLysIleThrTyrPheAlaLeu 156

Db 243 GAAGGAGACTTTTGGCGAGATCAAGCATGTTGGGCTCAAGATCACCTACTTTGCACTG 302
Qy 157 MetAspGlyLysValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSer 176
Db 303 ATGGATGGAAGGTGTATGACATCACAGAGTGGGCTGGATGCCAGCGTGTAGTATCTCC 362
Qy 177 ProAspThrHisArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArg 196
Db 363 CCAGATACCCACAGAGTCCCTATCAATCTCATTTGGTTCTCGATTCAGGACCAGAGA 422
Qy 197 GlyArgGlnArgAlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArg 216
Db 423 GGGCGGCGAGAGGCCACCCAGATGCCCCCTCTCTGCTGATCTTCAGGATTTCTTGAGTCGG 482
Qy 217 IlePheGlnValProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnPro 236
Db 483 ATCTTTCAAGTACCCCGGAGGAGATGCCCAATGCGAACTTCTTTGAGCTCCTCAGCCT 542
Qy 237 AlaProGlyAlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLys 256
Db 543 GCCCTGGAGCGCTGCAGCGCTTAAGCCCAACAGCAGTACCACAAAGGAGGAGGCCAAA 602
Qy 257 ProLysArgArgLysLysValArgArgProPheGlnArg 269
Db 603 CCTAAGCGCGGAAGAAAGTGGAGGCGCTTCCAAAGT 641

RESULT 14
BD020639 747 bp DNA linear PAT 27-AUG-2002
LOCUS Novel gene and novel gene fragment cloned in human neuroblastoma.
DEFINITION BD020639
ACCESSION BD020639
VERSION BD020639.1 GI:22561815
KEYWORDS JP 2001245671-A/2877.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 747)
AUTHORS Nakagawara,A.
TITLE Novel gene and novel gene fragment cloned in human neuroblastoma
JOURNAL Patent: JP 2001245671-A 2877 11-SEP-2001;
CHIBA PREF,HISAMITSU PHARMACEUTICAL CO INC

COMMENT OS Homo sapiens (human)
PN JP 2001245671-A/2877
PD 11-SEP-2001
PF 07-MAR-2000 JP 2000159195
PI AKIRA NAKAGAWARA

PC C12N15/09,C12Q1/68,G01N33/53,G01N33/566/(C12Q1/68,C12R1:91),
PC C12N15/00
CC Novel gene and novel gene fragment cloned in human CC
neuroblastoma
FH Key Location/Qualifiers
FT source 1..747
FT /organism='Homo sapiens (human)'.

FEATURES
source 1..747
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

BASE COUNT 191 a 167 c 200 g 165 t 24 others
ORIGIN

Alignment Scores:
Pred. No.: 2,36e-69 Length: 747
Score: 1112.00 Matches: 216
Conservative: 2
Percent Similarity: 91.21% Mismatches: 19
Best Local Similarity: 90.38% Indels: 3
Query Match: 77.60% Gaps: 1
DB: 6

US-10-049-742-11 (1-269) x BD020639 (1-747)

Qy 15 LeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAla 34

Db 36 TTGGGCACCTGTTGGCCTACTGGCAGATGTTGAACGTGAAGAGGCCTATAGACAGCTGGCA 95
Qy 35 ValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPheLysValLeu 54
Db 96 GTGATGGTTTCATCTCTGACAAAATCATCATCCCGGGCTGAGGAGCGCTTCAAGGTTTG 155
Qy 55 ArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMetLysArg 74
Db 156 CGAGCAGCTTGGGACATTTGTACAAATGCTGAAAAGCGAAAGGAGTATGATGAAACGA 215
Qy 75 MetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeuGlnAspAsp 94
Db 216 ATGGCAGAGAATGAGCTGAGCCGCTCAGTAAATGAGTTTCTGTCCAAGCTGCAAGATGAC 275
Qy 95 LeuLysGluAlaMetAsnThrMetMetCysSerArgCysGlnGlyLysHisArgArgPhe 114
Db 276 CTCAAGGAGGCAATGAATACTATGATGTGTAGCGGATGCCAAGGAAAGCATAGGAGGTTT 335
Qy 115 GluMetAspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisPro 134
Db 336 GAAATGACCGGGAACCTTAAGAGTGCAGATACGTGCTGAGTGTATAGGCTGCATCCT 395
Qy 135 AlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPhe 154
Db 396 GCTGAGGAGGAGACITTTTGGCAGAGTCAAGCATGTTGGGCTCAAGATCACCTACTTT 455
Qy 155 AlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAlaGlyCysGln-ArgValG1 174
Db 456 GCATGATGATGAAAGGTGATGACATCACAGATGGGCTGGATGCCACCGCTGTAGG 515
Qy 174 YlleSerProAspThrHisArgValProTyrHisIleSerPheGlySerArgIleProG1 194
Db 516 TATCTCCACAGATACCCACAGAGTCCCTCATCATCTTGGTTCTCGGATTCACAG 575
Qy 194 YThrArgGlyArgGlnArgAlaThrProAspAlaProProAlaAspLeuGlnAspPheLe 214
Db 576 CACCAGAGGGCGGAGAGAGCCACCCAGATGCCCCCTTCTGCTGATCTTCAGGATTTCTT 635
Qy 214 uSerArgIlePheGlnValProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaPr 234
Db 636 GAGTCGAATCTTCAAGTACCCCGGAGGCAATGCCCAATGGGAATCTTTCAGGATTC 695
Qy 234 oGlnProAlaProGlyAlaAlaAlaAlaSerLysProAsnSerThrValProLys 252
Db 696 TAA-CTTGGCCCTGGACCCGTTGAACNTTAAGC---CAACAGNACATACCCAAG 746

RESULT 15
BD100577 747 bp DNA linear PAT 27-AUG-2002
LOCUS Novel genes cloned in humanneuroblastoma and fragments thereof.
DEFINITION BD100577
ACCESSION BD100577.1 GI:22646151
VERSION WO 0166719-A/2877.
KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 747)
AUTHORS Nakagawara,A.
TITLE Novel genes cloned in humanneuroblastoma and fragments thereof
JOURNAL Patent: WO 0166719-A 2877 13-SEP-2001;
CHIBA PREF,HISAMITSU PHARMACEUTICAL CO INC,AKIRA NAKAGAWARA

COMMENT OS Homo sapiens (human)
PN WO 0166719-A/2877
PD 13-SEP-2001
PF 02-MAR-2001 WO 2001JP001629
PR 07-MAR-2000 JP 00P 159195
PI AKIRA NAKAGAWARA
PC C12N15/11,C12Q1/68,G01N33/53,G01N33/566
CC Novel genes cloned in humanneuroblastoma and fragments thereof
FH Key Location/Qualifiers
FT source 1..747

FT /organism='Homo sapiens (human)'

FEATURES
source

Location/Qualifiers
1..747
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 191 a 167 c 200 g 165 t 24 others
ORIGIN

Alignment Scores:

Pred. No.:	2,36e-69	Length:	747
Score:	1112.00	Matches:	216
Percent Similarity:	91.21%	Conservative:	2
Best Local Similarity:	90.38%	Mismatches:	19
Query Match:	77.60%	Indels:	3
DB:	6	Gaps:	1

US-10-049-742-11 (1-269) x BD100577 (1-747)

QY	15	LeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAla	34
DB	36	TTGGGCACCTGTGGCCCTACTGGCAGATGTTGAAGGAGCCCTATAGACAGCTGGCA	95
QY	35	ValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPheLysValLeu	54
DB	96	GTGATGGTTCATCTGCACAAAATCATCATCCCGGGCTGAGGAGCCCTCAAGGTTTG	155
QY	55	ArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMetLysArg	74
DB	156	CGAGCAGCTGGGACATTTGCAGCAATGCTGAAAAGCGAAAGGAGTATGAGATGAAACGA	215
QY	75	MetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLysLeuGlnAspAsp	94
DB	216	ATGGCAGAGAATGAGCTGACCGCTCAGTAAATGAGTTCTGTCCAAGCTGCAAGATGAC	275
QY	95	LeuLysGluAlaMetAsnThrMetCysSerArgCysGlnGlyLysHisArgArgPhe	114
DB	276	CTCAAGGAGCAATGAATACTATGATGTGTAGCGGATGCCAAGGAAGCATAGAGGTTT	335
QY	115	GluMetAspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisPro	134
DB	336	GAATGGACCGGAACCTAAGAGTGCAGACTACTGTGCTGAGTGTAAATAGCTGCATCCT	395
QY	135	AlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPhe	154
DB	396	GCTGAGGAAGGAGACTTTTGGGCAGAGTCAAGCATGTGGCCCTCAAGATCACCTACTTT	455
QY	155	AlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAlaGlyCysGln-ArgValG1	174
DB	456	GCACATGATGGATGAAAGGTGTATGACATCACAGAGTGGGCTGGATGCCAGCCGTGTAGG	515
QY	174	YlleSerProAspThrHisArgValProTyrHisIleSerPheGlySerArgIleProG1	194
DB	516	TATCTCCCCAGATACCCACAGAGTCCCTCATCATCTCATTTGGTTCTCGGATTCAGG	575
QY	194	YThrArgGlyArgGlnArgAlaThrProAspAlaProProAlaAspLeuGlnAspPheLe	214
DB	576	CACGAGGGCGGCGAGAGCCACCCAGATGCCCTTCTGCTGATCTTCAGGATTTCTT	635
QY	214	uSerArgIlePheGlnValProProGlyGlnMetProAsnGlyAsnPheAlaAlaPr	234
DB	636	GAGTCGAATCTTCAAGTACCCCGGCGCANATGCCCAATGGGAACITCTTTGCGAGCTC	695
QY	234	oGlnProAlaProGlyAlaAlaAlaAlaSerLysProAsnSerThrValProLys	252
DB	696	TAA-CCTTGGCCTGGACCCGTTGAANCTNTAAGC---CAACAGNACATACCCCAAG	746

Search completed: December 1, 2003, 11:25:05
Job time : 4169 secs

and is derived by analysis of the total score distribution.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - nucleic search, using frame_plus_p2n model

Run on: December 1, 2003, 07:37:22 ; Search time 318 Seconds
(without alignments)
2283.487 Million cell updates/sec

Title: US-10-049-742-11
Perfect score: 1433
Sequence: 1 MAGVPEDELNPFHLVGEAT.....VPGKAKPKRRKVRPPQR 269

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2522756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODE=frame+p2n.model -US10049742/runat_01122003_073135_18565/app_query.fasta_1.455
-O=/cgr2_1/USPTO_spool/US10049742 -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-DB=N_Geneseq_19Jun03 -QFMT=fastap -MATRX=blomsum62 -TRANS=human40.cdi
-LCOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10049742 @CGN_1_1_312/runat_01122003_073135_18565 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1433	100.0	1147	22	AAF54994	Nucleotide sequenc
2	1430	99.8	2279	24	ABL55088	Human cDNA sequenc
3	1430	99.8	3286	24	ABL56700	Nucleotide sequenc
4	1112	77.6	747	22	AAI96802	Human neuroblastom
5	1051	73.3	754	20	AAZ34492	Human LYST interac
6	1051	73.3	754	24	ABX04184	Human mRNA differe
7	1051	73.3	754	24	ABX76460	cDNA encoding huma
8	1012.5	70.7	7453	22	AA330481	DNA encoding novel
9	1012.5	70.7	7453	22	AA330481	Human reproductive
10	497	34.7	12119	22	AA330481	Nucleotide sequenc
11	497	34.7	12578	21	AA330481	Nucleotide sequenc
12	497	34.7	12734	24	ABA95615	Chimeric BVDV/HCV
13	497	34.7	12734	24	ABA95615	Nucleotide sequenc
14	497	34.7	12842	21	AAZ36203	Nucleotide sequenc
15	497	34.7	13198	21	AAZ36211	Plasmid pBVDN1, c
16	497	34.7	14078	21	AA338807	Bovine viral diar
17	497	34.7	14578	21	AAZ48136	Nucleotide sequenc
18	497	34.7	15065	21	AAZ36195	Nucleotide sequenc
19	469	32.7	11674	21	AAZ36210	Nucleotide sequenc
20	416	29.0	3135	23	ABL25653	Drosophila melanog
21	329	23.0	5447	23	ABL25652	Drosophila melanog
22	321	22.4	884	24	ABL55114	Human cDNA sequenc
23	211	14.7	330	16	AAI24617	Human gene signatu
24	195.5	13.6	1030	22	AAZ27140	cDNA encoding nove
25	195.5	13.6	1030	22	AAZ27140	cDNA encoding nove
26	189	13.2	1283	22	AAH79328	Human DnaJ protein
27	189	13.2	1307	24	ABN59622	Novel human coding
28	189	13.2	1764	21	AAA08564	Human chaperone pr
29	183	12.8	378	22	AAF65696	Novel human polynu
30	182	12.7	1413	23	ABL05587	Drosophila melanog
31	179	12.5	1645	22	AAZ27141	cDNA encoding nove
32	173	12.1	323	22	AAI14193	Human breast cance
33	173	12.1	323	22	AAI14193	Human breast cance
34	173	12.1	4103	23	ABL05586	Drosophila melanog
35	172.5	12.0	1993	25	ABX55986	Gene encoding huma
36	172.5	12.0	2239	23	ABV24858	Human prostate exp
37	172.5	12.0	3001	24	ABZ35537	Human gene express
38	172	12.0	1046	22	AAZ27514	cDNA encoding nove
39	171	11.9	2918	25	ABT13422	Breast specific re
40	170	11.9	1430	22	AA331094	Human diagnostic a
41	169	11.8	767	22	AAH07702	Human cDNA clone (
42	169	11.8	2303	22	AAH18620	Human cDNA sequenc
43	165	11.5	501	22	AAH06260	Human cDNA clone (
44	165	11.5	650	22	AAZ27511	cDNA encoding nove
45	165	11.5	993	22	AAH14366	Human cDNA sequenc

ALIGNMENTS

RESULT 1

AAF54994

ID AAF54994 standard; DNA; 1147 BP.

XX AAF54994;

AC AAF54994;

XX 15-MAY-2001 (first entry)

DT Nucleotide sequence of a human chaperone polypeptide.

DE Human; chaperone polypeptide; reproductive disease; prolactin production;

XX infertility; tumour; cancer; Peyronie's disease; eye disorder; glaucoma;

KW conjunctivitis; keratitis; neuromuscular disorder; cystic fibrosis;

KW metabolic disorder; Zellweger syndrome; Addison's disease; iritis;

KW autoimmune disorder; inflammatory disorder; systemic lupus erythematosus;

KW acquired immunodeficiency syndrome; AIDS; aschma; atherosclerosis;

KW cell proliferative disorder; gene therapy; ss.

XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 98..907 /*tag= a
 FT FT /product= "chaperone protein"
 FT FT
 XX WO200109178-A2.
 XX PN
 XX PD
 XX PD 08-FEB-2001.
 XX PF 03-AUG-2000; 2000WO-US21313.
 XX PR 03-AUG-1999; 99US-0146908.
 XX PR 22-OCT-1999; 99US-0160924.
 XX PA (INCY-) INCYTE GENOMICS INC.
 XX XX
 XX PI Yue H, Bandman O, Tang YT, Baughn MR, Azimzai Y, Lu DAM;
 XX DR P-PSDB; AAB67455.
 XX DR
 XX XX New human chaperone proteins and polynucleotides, useful in diagnosing,
 PT treating and preventing reproductive, eye, neuromuscular, metabolic,
 PT autoimmune or inflammatory disorders
 XX XX
 XX PS Claim 5; Page 95-96; 102pp; English.
 XX CC The present sequence encodes a human chaperone polypeptide. Human
 CC chaperone polypeptides and polynucleotides are useful in the diagnosis,
 CC treatment and prevention of reproductive (e.g. prolactin production,
 CC infertility, endometrial or ovarian tumour, cancer of the breast,
 CC prostate or testis, Peyronie's disease), eye (e.g. conjunctivitis,
 CC keratitis, iritis, glaucoma), neuromuscular, metabolic (e.g. Zellweger
 CC syndrome, Addison's disease, cystic fibrosis), and autoimmune and
 CC inflammatory disorders (e.g. systemic lupus erythematosus, acquired
 CC immunodeficiency syndrome (AIDS), asthma, atherosclerosis), infectious
 CC or viral diseases, and cell proliferative disorders. Chaperone
 CC polynucleotides may be used for somatic or germline gene therapy, to
 CC detect and quantify gene expression in biopsied tissues in which
 CC expression is correlated with disease.
 XX XX
 XX SQ Sequence 1147 BP; 287 A; 278 C; 314 G; 268 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.4e-116 Length: 1147
 Score: 1433.00 Matches: 269
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-10-049-742-11 (1-269) x AAF54994 (1-1147)
 QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
 DB 98 ATGGCTGGGGTCTCTGAGGATGAGCTTAACCTTTCCATGTTACTGGGGTGTAGGCCACA 157
 QY 21 AlaSerAspValGluLeuLysLysAlaTyArgGlnLeuAlaValMetValHisProAsp 40
 DB 158 GCATCAGATGTTGACTGAAGAGGCCCTATAGACAGCTGGCAGTGATGGTTTCATCCTGAC 217
 QY 41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
 DB 218 AAAAATCATCATCCCCGGGTGAGGAGGCCCTCAAGGTTTTCGAGAGCGCTTGGGACATT 277
 QY 61 ValSerAsnAlaGluLysArgLysGluTyGluMetLysArgMetAlaGluAsnGluLeu 80
 DB 278 GTCAGCAATCTGAAAAGCCGAGGAGTATGAGATGAACCAAGATGGCAGAGAAATGAGCTG 337
 QY 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100

DB 338 AGCCGGTCAGTAAATGAGTTTCTGTCCAAGTGAAGATGACCTCAAGAGGCAATGAT 397
 QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120
 DB 398 ACTATGATGTGTAGCCGATGCCAAGAAAGATAGGAGTTTGAATGACCGGAACT 457
 QY 121 LysSerAlaArgTyCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
 DB 458 AAGAGTGCCAGATACTGTGCTGAGTGTAAATAGGTGTCATCTCTGTGAGAGGAGACTTT 517
 QY 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyPheAlaLeuMetAspGlyLys 160
 DB 518 TGGGAGAGTCAAGCATGTTGGGCTCAAGATCACTACTTTGCACTGATGATGAAAG 577
 QY 161 ValTyAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
 DB 578 GTGTATGACATCACAGATGGGCTGGATGCCAGCGGTAGGTATCTCCCAATATACCCAC 637
 QY 181 ArgValProTyHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
 DB 638 AGAGTCCCCTATCATCTCTCTTGGTTCTCGATTCCAGGCCACAGAGGCGGCAGAGA 697
 QY 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
 DB 698 GCCACCCAGATGCCCTCTCTGCTGATCTTCAAGATTCTTCAAGTGGATCTTTCAAGTA 757
 QY 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
 DB 758 CCCCAGGCGAGATGCCCAATCGGAACCTTCTTTGAGCTCTCCAGCTCCCTCGAGCC 817
 QY 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyValAlaLysProLysArgArg 260
 DB 818 GCTGAGCGCTCTAAGCCCAACAGCAGTACCACCAAGGAGAAAGCCAACTTAAGCGCGG 877
 QY 261 LysLysValArgArgProPheGlnArg 269
 DB 878 AAGAAAGTGAGGAGGCCCTTCCAAAGT 904
 RESULT 2
 ID ABL55088
 AB ABL55088 standard; DNA; 2279 BP.
 XX AC ABL55088;
 XX XX
 DT 08-OCT-2002 (first entry)
 XX XX
 DE Human cDNA sequence #1 from clone HNTPB82.
 XX XX
 KW Human; HNTPB82; secreted protein; immunosuppressive; food preservative;
 KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac;
 KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
 KW virucide; fungicide; ophthalmological; vulnery; gene therapy; ELISA;
 KW radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;
 KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
 KW cerebrovascular disorder; nervous system disorder; ocular disorder;
 KW wound healing; food additive; ss.
 XX XX
 OS Homo sapiens.
 XX Key Location/Qualifiers
 XX CDS 36..1274 /*tag= a
 FT sig_peptide 36..38 /*tag= b
 FT mat_peptide 39..1271 /*tag= c
 XX WO200222638-A1.
 XX XX
 XX PD 21-MAR-2002.
 XX XX
 XX PF 17-JAN-2001; 2001WO-US01386.

XX 12-SEP-2000; 2000US-232104P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CS, Soppet DR, Olsen HS;
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI Ni J;
XX WPI; 2002-258041/30.
DR ?-PSDB; ABB77019.
XX
XX New nucleic acid molecules encoding 22 human secreted proteins for
PT diagnosing or treating e.g. autoimmune diseases, hyperproliferative
PT disorders, and cardiovascular disorders, and used as food additives or
PT preservatives -
XX
XX Disclosure; Page 451-452; 526pp; English.
XX
XX The sequence represents a cDNA sequence of the invention, isolated
CC from human clone ID HNTB82. The invention relates to novel
CC isolated nucleic acid molecules encoding 22 human secreted proteins. The
CC proteins of the invention have immunosuppressive, antiarthritic,
CC antirheumatic, antiproliferative, cytostatic, cardiant, vasotropic,
CC cerebroprotective, neurotropic, neuroprotective, antibacterial, virucide,
CC fungicide, ophthalmological, and vulnerary activity. The polynucleotides
CC may have a use in gene therapy. The polynucleotides and polypeptides
CC encoded by them are used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. The polynucleotides and polypeptides are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to the proteins and in diagnostic immunoassays
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases,
CC hyperproliferative disorders, cardiovascular disorders, cerebrovascular
CC disorders, angioneurosis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The polypeptides can
CC also be used to aid wound healing and epithelial cell proliferation. The
CC polypeptides can also be used as a food additive or preservative.
XX
XX Sequence 2279 BP; 540 A; 518 C; 607 G; 614 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 1,466-115 Length: 2279
Score: 1430.00 Matches: 268
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.63% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 24 Gaps: 0

US-10-049-742-11 (1-269) x ABL55088 (1-2279)

Qy 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
Db 465 ATGGCTGGGGTTCCTGAGGATGAGTAAACCCCTTCCATGTACTGGGGTTGAGGCCACA 524
Qy 21 AlaSerAspValGluLeuLysLysAlaTyArgGlnLeuAlaValMetValHisProAsp 40
Db 525 GCATCAGATGTTGAACGTAAGAGGCGCTATAGACAGCTGGCATGTGTTTCATCTGAC 584
Qy 41 LysAsnHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
Db 585 AAAAAATCATATCCCCGGCTGAGGAGGCTTCAAGGTTTTCGACAGCTGGGACATT 644
Qy 61 ValSerAsnAlaGluLysArgLysGluTyArgMetLysArgMetAlaGluAsnGluLeu 80
Db 645 GTCAGCAATGCTGAAAGCGAAGGAGTATGACATGAACGAATGCAGAGAAATGAGCTG 704
Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLysGluAlaMetAsn 100
Db 705 AGCCGGTCAGTAAATGAGTTTCTGTCCAAAGCTCAAGATGACCTCAAGGAGGCAATGAAT 764

Qy 101 ThrMetMetCysSerArgCysGlnGlyHisArgArgPheGluMetAspArgGluPro 120
Db 765 ACTATGATGTGTAGCCGATGCCAAGGAAAGCATAGGAGGTTTGAAATGGACCGGAACCT 824
Qy 121 LysSerAlaArgTyArgCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
Db 825 AAGAGTGCCAGATACTGTGCTGAGTGAATAGGCTGATCTCTGCTGAGGAGGAGACATT 884
Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyPheAlaLeuMetAspGlyLys 160
Db 885 TGGCAGAGTCACAGCATGTTGGGCTCAAGATCACCTACTTTGCACATGATGATGAAAG 944
Qy 161 ValTyAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
Db 945 GTGTATGACATCACACATGGGCTGGATGCCAGGCTGTAGGTATCTCCCCAGATACCCAC 1004
Qy 181 ArgValProTyHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
Db 1005 AGAGTCCCTATCATCTCATTTGGTTCTCGGATTCAGGACCAGAGGGCGGAGAGA 1064
Qy 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
Db 1065 GCCACCCAGATGCCCTCTCTGCTGATCTTCAAGATTTCTTGAGTCGGATCTTTCAAGTA 1124
Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
Db 1125 CCCCAGGCGCAGATGCCCAATGGGAACCTTTTGCAGCTCTCAGCTGCCCTGGAGCC 1184
Qy 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
Db 1185 GCTGAGGCTCTAAGCCCAACAGCAGCAGTACCAAGGAGGAGGAGCAAACTAAGCGCGG 1244
Qy 261 LysLysValArgArgProPheGlnArg 269
Db 1245 AAGAAAGTGAGGAGGCCCTTCCACGT 1271
RESULT 3
ABU56700
ID ABL56700 standard; cDNA; 3286 BP.
XX
XX ABL56700;
XX
XX 30-JUL-2002 (first entry)
XX
XX Nucleotide sequence of human P125-77.22 polypeptide.
XX Human; P125-77.22; mucosal disease; BVDV infection; gene therapy;
XX gene; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 122..2230
XX FT /*tag= a
XX FT /product= "P125-77.22 polypeptide"
XX PN W020226810-A1.
XX
XX 04-APR-2002.
XX PD
XX PF 10-SEP-2001; 2001WO-CN01354.
XX PR 12-SEP-2000; 2000CN-0125190.
XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX
XX Mao Y, Xie Y;
XX PI
XX WPI: 2002-281319/32.
XX P-PSDB; ABB77732.
XX
XX Protein P125-77.22 and encoding polynucleotide, used in diagnosis and
PT treatment of human mucosal disease caused by BVDV infection -
PT

XX PS Claim 6; Page 27-29; 33pp; Chinese.
XX CC The present sequence encodes human P125-77-22 polypeptide. The
CC polypeptide and polynucleotide are used in diagnosis and treatment
CC of human mucosal disease caused by viral BVDV (undefined) infection.
CC The polynucleotide may also be used for gene therapy.
XX SQ Sequence 3286 BP; 787 A; 785 C; 895 G; 819 T; 0 other;

Alignment Scores:
Pred. No.: 2,3e-115 Length: 3286
Score: 1430.00 Matches: 268
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.63% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 24 Gaps: 0

US-10-049-742-11 (1-269) x ABL56700 (1-3286)
Qy 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
Db 1421 ATGGCTGGGGTTCTCTGAGGATGAGCTAAACCTTTCCATGTACTGGGGTTGAGGCCACA 1480
Qy 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
Db 1481 GCATCAGATGTTGAACCTGAGGAGGCTATACACAGCTGGCAGTGTATCTCTGAC 1540
Qy 41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
Db 1541 AAAATATCATCTCCCGGGCTGAGGAGGCTTCAGGTTTTCGAGCAGCTGGGACATT 1600
Qy 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
Db 1601 GTACAGCAATGCTGAAAGCGGAAAGGAGTATGAGATGAAACGAATGCGAGAAATGAGCTG 1660
Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsn 100
Db 1661 AGCGGTGCTAGTAATGAGTTCTGTCTCAAGCTGCAAGATGACCTCAAGGAGCAATGAAT 1720
Qy 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120
Db 1721 ACTATGATGTGTAGCGATGCCAAGGAAAGCATAGGAGGTTTGAAATGACCGGGAACCT 1780
Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
Db 1781 AAGAGTGCTAGATGCTGTGAGTGTATGAGTGTGCTGCTGCTGAGGAGGAGACTTT 1840
Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 1841 TGGGCGAGGTCAAGCATGTTGGGCTCAAGATCACCTACTTTCACCTGATGGATGGAAG 1900
Qy 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyLysSerProAspThrHis 180
Db 1901 GTGTATGACATCACAGATGGGTGGATGCCAGCGTGTAGTATCTCCCGAGATACCCAC 1960
Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
Db 1961 AGAGTCCCCCTATCACATCTCATTTGGTTCTTCGAATCCAGGCACGAGGCGGCGAGAGA 2020
Qy 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
Db 2021 GCCACCCAGATGCCCTCTCTGCTGATCTTCAGGATTTCTTGAGTCGATCTTTCAAGTA 2080
Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
Db 2081 CCCCCAGGCGAGATGCCCAATGGGAATCTCTTTCAGCTCTCTCAGCTGCCCCCTGGAGCC 2140
Qy 241 AlaAlaLaserLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
Db 2141 GCTGACGCTCTAAGCCCAACAGCACAGTACCCCAAGGGAGGAAGCAACCTTAAGCGCGCG 2200
Qy 261 LysLysValArgArgProPheGlnArg 269

Db 2201 AAGAAAGTGAGGAGGCCCTTCCAAACGT 2227
RESULT 4
AAI96802
ID AAI96802 standard; cDNA; 747 BP.
XX AC AAI96802;
XX DT 13-NOV-2001 (first entry)
XX DE Human neuroblastoma expressed polynucleotide SEQ ID NO 2877.
XX KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX OS Homo sapiens.
XX PN WO200166719-A1.
XX PD 13-SEP-2001.
XX PF 02-MAR-2001; 2001WO-JP01629.
XX PR 07-MAR-2000; 2000JP-0159195.
XX PA (CHIB-) CHIBA PREFECTURE.
XX PA (HISM-) HISAMITSU PHARM CO LTD.
XX PI Nakagawara A;
XX WIPI; 2001-565584/63.
XX PT Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT for anti-cancer agents -
XX PS Claim 1; Page 2103; 2979pp; Japanese.
XX CC The invention relates to novel genes (AAI93926-AAI97963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes.
XX SQ Sequence 747 BP; 191 A; 167 C; 200 G; 165 T; 24 other;

Alignment Scores:
Pred. No.: 2,51e-88 Length: 747
Score: 1112.00 Matches: 216
Percent Similarity: 91.21% Conservative: 2
Best Local Similarity: 90.38% Mismatches: 19
Query Match: 77.60% Indels: 3
DB: 22 Gaps: 1

US-10-049-742-11 (1-269) x AAI96802 (1-747)
Qy 15 LeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAla 34
Db 36 TTGGGCACCTGTGGCCCTACTGCGAGATGTTGAATCAAGAGGCTTATAGACAGCTGGCA 95
Qy 35 ValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPheLysValLeu 54
Db 96 GTGATGTTTCATCTCCACAAAATCATCTCCCGGCTGAGGAGGCTTCAAGGTTTG 155
Qy 55 ArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMetLysArg 74
Db 156 CGAGCAGCTGGGACATTTGTCAGCAATGCTGAAAAGCGAAGAGGATATGAGATGAACGA 215
Qy 75 MetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeuGlnAspAsp 94
Db 216 ATGCGCAGAGAAATGAGCTGAGCGGCTCAGTAAATGAGTTTCTGTCCCAAGCTGCAAGATGAC 275

QY 95 LeuLysGluAlaMetAsnThrMetCysSerArgCysGluGlyHisArgArgPhe 114
DB 276 CTTAAGGAGCGCAATGAATCTATGATGTAGTCCGATGCCAAGGAAGCATAGGAGTTT 335
QY 115 GluMetAspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisPro 134
DB 336 GAAATGGACCGGAACCTAAGAGTGCAGATACTGTCTGAGTGTAATAGGCTGCATCT 395
QY 135 AlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPhe 154
DB 396 GCTGAGGAAGGAGACTTTTGGGCGAGAGTCAAGCATGTTGGGCTCAAGATCACCTACTTT 455
QY 155 AlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAlaGlyCysGlnArgValG 174
DB 456 GCACCTGATGGATGAAGAGTGTATGACATCACAGAGTGGGCTGGATGCCAGCGGTGAG 515
QY 174 YleSerProAspThrHisArgValProTyrHisIleSerPheGlySerArgIleProG 194
DB 516 TATCTCCCCAGATACCCACAGAGTCCCTATACATCTCTATTGGTTCGGATTCAGG 575
QY 194 YThrArgGlyArgGlnArgAlaThrProAspAlaProProAlaAspLeuGlnAspPheLe 214
DB 576 CACCAGAGGGCGCAGAGAGCCACCCAGATGCCCTTCTGCTGATCTTCAGGATTTCTT 635
QY 214 uSerArgIlePheGlnValProGlyGlnMetProGlyGlnMetProAsnGlyAsnPheAlaAlaPr 234
DB 636 GAGTGAATCTTCAAGTAGTACCCCGGCGCANATGCCAATGGGAATCTTTTGCAGCTTC 695
QY 234 cGlnProAlaProGlyAlaAlaAlaSerLysProAsnSerThrValProLys 252
DB 696 TAA-CCTTCCCTGGACCCGTTTGAANCTNTAAGC---CAACAGNACATACCCCAAG 746
RESULT 5
ID AA234492 standard; cDNA; 754 BP.
XX
AC AA234492;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human LYST interacting protein LIP6 cDNA.
XX
KW LIP6; human; LYST-2; LYST interacting protein; lysosome; CHS;
KW Chediak-Higashi syndrome; autoimmune disease; rheumatoid arthritis;
KW systemic lupus erythematosus; inflammatory bowel disease;
KW diabetes mellitus; multiple sclerosis; atopic disease; asthma;
KW hay fever; rhinitis; urticaria; nasal polyp; cancer;
KW neurodegenerative disease; pigmentation disorder; viral disease;
KW platelet dysfunction; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 11..586
FT /*tag= a
FT /*partial
XX
PN WO9951741-A2.
XX
PD 14-OCT-1999.
XX
PF 29-MAR-1999; 99WO-US06831.
XX
PR 03-APR-1998; 98US-0054956.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Nandabalan K, Kingsmore S;
XX
DR WPI; 1999-620203/53.
DR P-PSDB; AAY32126.
XX

PT Protein complexes, interacting proteins, and related polynucleotides
PT useful for treating and preventing e.g. atopic, autoimmune or
PT neurodegenerative diseases -
XX
PS Claim 21; Fig 7; 172pp; English.
XX
CC This is the nucleotide sequence of cDNA which codes for a novel
CC human LYST interacting protein, LIP6 (see AAY32126), that shows
CC homology to pestivirus NS2-3. LYST is the human lysosomal
CC Chediak-Higashi syndrome (CHS) protein. The invention relates to
CC complexes of LYST or LYST-2 (see AAY32120) with proteins identified
CC as interacting with LYST or LYST-2 by a modified yeast two-hybrid
CC assay system. The interacting proteins include 10 novel proteins,
CC LIP1-10 (see AAY32121-30). Methods of screening the protein complexes
CC for efficacy in treating and/or preventing atopic diseases (e.g.
CC asthma, nasal polyps, hay fever rhinitis, urticaria) autoimmune
CC diseases (e.g. CHS, rheumatoid arthritis, systemic lupus
CC erythematosus, inflammatory bowel disease, diabetes mellitus,
CC multiple sclerosis), neurodegenerative disease, certain forms
CC of cancer, pigmentation disorders, platelet dysfunction and viral
CC diseases are provided. Nucleic acids (see AA234487-96) encoding
CC LIP1-10, modulation of LIP function by gene therapy, use of
CC antisense oligonucleotides for suppression of LIP protein
CC expression, screening for agonists and antagonists, diagnosing or
CC screening for the presence of a predisposition to a disease or
CC disorder, and animal models are also disclosed.
XX
SQ Sequence 754 BP; 193 A; 185 C; 200 G; 176 T; 0 other;

Alignment Scores:
Pred. No.: 5,55e-83 Length: 754
Score: 1051.00 Matches: 200
Percent Similarity: 99.50% Conservative: 0
Best Local Similarity: 99.50% Mismatches: 1
Query Match: 73.34% Indels: 1
DB: 20 Gaps: 0

US-10-049-742-11 (1-269) x AA234492 (1-754)

QY 69 GluTyrGluMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeu 88
DB 2 GAGTATGAGTGAACCAATGCGCAGAAATGAGTCCGCGGTCAGTAAATGATTTCTG 61
QY 89 SerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMetCysSerArgCysGln 108
DB 62 TCAAGCTGCAAGATGACCTCAAGAGGCAATGAATACTATGATGTAGCCCATGCCAA 121
QY 109 GlyLysHisArgArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGlu 128
DB 122 GGAAGCATAGGAGGTTTGAATGGACCCGGAACCTAAGAGTCCAGATACTGTCTGAG 181
QY 129 CysAsnArgLeuHisProAlaGluGlyAspPheTrpAlaGluSerSerMetLeuGly 148
DB 182 TGTATAGGCTGCATCCTGCTGAGGAAGGAGACTTTTGGGCGAGTCAAGCATGTTGGGC 241
QY 149 LeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAla 168
DB 242 CTCAGATCACCTACTTTGCTGATGGATGGAAGGTGTATGACATCACAGATGGGCT 301
QY 169 GlyCysGlnArgValGlyIleSerProAspThrHisArgValProTyrHisIleSerPhe 188
DB 302 GGATGCCAGCGTGTAGGTATCTCCCGAGATACCCAGAGTCCCTTATCATCTCATTT 361
QY 189 GlySerArgIleProGlyThrArgGlyArgGlnArgAlaThrProAspAlaProAla 208
DB 362 GGTTCCTCGGATTCAGGCACAGAGCGGCGCAGAGAGCCACCCAGATGCCCTCTCTGCT 421
QY 209 AspLeuGlnAspPheLeuSerArgIlePheGlnValProGlyGlnMetProAsnGly 228
DB 422 GATCTTCAGGATTTCTTGAGTCGGATCTTCAAGTACCCCGGCGAGATGCCAAT-GGG 480
QY 229 AsnPhePheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerLysProAsnSer 248

Db 481 AACTTCTTTGAGCTCTCAGCCTGCGCCCTGGAGCGCTGAGCGCTCTTAAGCCCAACAGC 540
 Qy 249 ThrValProLysGlyGluAlaLysProLysArgArgLysLysValArgArgProPheGln 268
 Db 541 ACAGTACCCAAAGGAGGAAGCAACCTTAAGCGGCGGAAGAAAGTGAGGAGCGCTTCCAA 600
 Qy 269 Arg 269
 Db 601 CGT 603
 RESULT 6
 ID ABX04184
 XX ABX04184 standard; cDNA; 754 BP.
 AC ABX04184;
 XX
 DT 10-JAN-2003 (first entry)
 XX Human mRNA differentially expressed in mesenchymal cells #31.
 XX Human; ss; gene; skeletal growth; cartilage degeneration disorder;
 KW chondroblastic phenotype; mesenchymal cell; cartilage formation;
 KW bone formation; arthritis; osteoarthritis; rheumatoid arthritis;
 KW gout arthritis; adjuvant arthritis; arthritis deformans; antigout;
 KW infectious arthritis; osteochondrosis; RDA; antiarthritic; osteopathic;
 KW antirheumatic; antinflammatory; representational difference analysis.
 XX
 OS Homo sapiens.
 XX
 EN W0200271927-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 12-MAR-2002; 2002WO-US07787.
 XX
 PR 12-MAR-2001; 2001US-274980P.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Yates KE, Mizuno S, Glowacki J;
 XX WPI; 2002-732276/78.
 DR
 XX New nucleic acid molecules capable of promoting chondrogenesis, useful
 PT for diagnosing and treating cartilaginous tissue degeneration
 PT conditions, e.g. osteoarthritis, rheumatoid arthritis, gout arthritis,
 PT or osteochondrosis -
 XX
 PS Claim 33; Page 129; 153pp; English.
 XX
 CC The invention relates to new isolated nucleic acid molecule comprising a
 CC nucleic acid molecule consisting of a gene differentially expressed
 CC in cells undergoing differentiation from mesenchymal cell to a
 CC chondroblastic phenotype, or hybridising under stringent conditions
 CC to them (or their fragments). Also included are expression vectors,
 CC transformed host cells, expressed polypeptides or peptide fragments
 CC (which induce differentiation of a mesenchymal cell and may be used as an
 CC immunogen), binding partners of the polypeptides, a method for
 CC identifying an agent useful in modulating mesenchymal cell
 CC differentiation induction activity of a molecule, a method of diagnosing
 CC a condition characterized by aberrant expression of a nucleic acid
 CC molecule or its expression product; a method for determining regression,
 CC progression or onset of cartilaginous tissue degeneration condition in a
 CC subject characterised by aberrant expression of a nucleic acid molecule
 CC or its expression product, a method for treating a cartilaginous tissue
 CC degeneration condition, a method for treating a subject to reduce the
 CC risk of cartilaginous tissue degeneration condition developing in the
 CC subject, a method for identifying a candidate agent for treating a
 CC cartilaginous tissue degeneration condition, and a solid-phase nucleic
 CC acid molecule array consisting essentially of a set of nucleic acid
 CC molecule as cited above (or known from known genes shown to be
 CC differentially expressed in developing mesenchymal cells using the
 CC technique of representational difference analysis, RDA), its expression

CC products or fragments, fixed to a solid substrate. The nucleic acids,
 CC polypeptides and agents are useful for treating cartilaginous tissue
 CC degeneration conditions such as osteoarthritis, rheumatoid arthritis,
 CC gout arthritis, adjuvant arthritis, arthritis deformans, infectious
 CC arthritis or osteochondrosis. The present sequence is a cDNA from
 CC a known gene differentially expressed in developing mesenchymal cells.
 XX

SQ Sequence 754 BP; 193 A; 185 C; 200 G; 176 T; 0 other;

Alignment Scores:

Pred. No.: 5,55e-83 Length: 754
 Score: 1051.00 Matches: 200
 Percent Similarity: 99.50% Conservative: 0
 Best Local Similarity: 99.50% Mismatches: 1
 Query Match: 73.34% Indels: 1
 DB: 24 Gaps: 0

US-10-049-742-11 (1-269) x ABX04184 (1-754)

Qy 69 GluTyrGluMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeu 88
 Db 2 GAGTATGAGATGAACCGAATGGCAGAGAATGAGCTGAGCGGTGAGTAAATGAGTTTCTG 61
 Qy 89 SerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGln 108
 Db 62 TCCAAGCTGCAAGATGACCTCAAGGAGCAATGAATACTATGATGTAGCCGATGCCAA 121
 Qy 109 GlyLysHisArgArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGlu 128
 Db 122 GGAAGCATAGGAGGTTTGAATGGACCGGGAACCTAAGAGTGCCACATATGCTGAG 181
 Qy 129 CysAsnArgLeuHisProAlaGluGlyAspPheTrpAlaGluSerSerMetLeuGly 148
 Db 182 TGTAATAGGCTGCATCCTGCTGAGGAAGGAGACTTTTGGGCAGAGTCAAGCATTTGGGC 241
 Qy 149 LeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAla 168
 Db 242 CTCAGATCACCTACTTTGCACTGATGGATGGAAGGTGATGACATCACAGATGGGCT 301
 Qy 169 GlyCysGlnArgValGlyIleSerProAspThrHisArgValProTyrHisIleSerPhe 188
 Db 302 GGATGCCAGCGGTAGGTATCTCCCGAGATACCCACAGATCCCTTATCATCTCATTT 361
 Qy 189 GlySerArgIleProGlyThrArgGlyArgGluArgAlaThrProAspAlaProAla 208
 Db 362 GGTCTCGGATTCAGGCACACAGAGGGCGGACAGAGACCCACCCAGATGCCCTCTGCT 421
 Qy 209 AspLeuGlnAspPheLeuSerArgIlePheGlnValProProGlyGlnMetProAsnGly 228
 Db 422 GATCTTCAGGATTTCTTGTAGTCGATCTTCAAGTACCCCGGCGAGATGCCAAT-GGG 480
 Qy 229 AsnPhePheAlaAlaProGlnProAlaProGlyAlaAlaAlaAlaSerLysProAsnSer 248
 Db 481 AACTTCTTTTCAGCTCCTCAGCTGCCCTGGAGCCCTGCAGCCTCTTAAGCCCAACAGC 540
 Qy 249 ThrValProLysGlyGluAlaLysProLysArgArgLysLysValArgArgProPheGln 268
 Db 541 ACAGTACCCAAAGGAGGAAGCAACCTTAAGCGGCGGAAGAAAGTGAGGAGCGCTTCCAA 600
 Qy 269 Arg 269
 Db 601 CGT 603
 RESULT 7
 ABS76460
 ID ABS76460 standard; cDNA; 754 BP.
 XX
 AC ABS76460;
 XX
 DT 11-DEC-2002 (first entry)
 XX
 DE cDNA encoding human ovarian cancer marker OV38.
 XX

KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker; gene; ss.
XX Homo sapiens.
OS

PN WO200271928-A2.
XX
XX 19-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US07826.
XX
XX 14-MAR-2001; 2001US-276025P.
PR 14-MAR-2001; 2001US-276026P.
PR 10-AUG-2001; 2001US-311732P.
PR 19-SEP-2001; 2001US-323580P.
PR 26-SEP-2001; 2001US-324967P.
PR 26-SEP-2001; 2001US-325102P.
PR 26-SEP-2001; 2001US-325149P.
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
XX WPI; 2002-723277/78.
DR P-PSDB; ABG96384.
XX

XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient -
XX
XX Disclosure; Page 300; 48ipp; English.

XX The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer,
CC determining whether ovarian cancer has metastasized or is likely to
CC metastasize, selecting a composition for inhibiting ovarian cancer,
CC assessing the ovarian carcinogenic potential of a compound, or
CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC present nucleic acid sequence encodes one of the ovarian cancer markers
CC described in the invention.

SQ Sequence 754 BP; 193 A; 185 C; 200 G; 176 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 5.55e-83 Length: 754
Score: 1051.00 Matches: 200
Percent Similarity: 99.50% Conservative: 0

Best Local Similarity: 99.50% Mismatches: 1
Query Match: 73.34% Indels: 1
DB: 24 Gaps: 0
US-10-049-742-11 (1-269) x ABS76460 (1-754)
QY 69 GluTyrGluMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeu 88
Db 2 GAGTATGAGATGAACGATGGCAGAGATGACCTGAGCCGGTCAGTAATGAGTTCTG 61
QY 89 SerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGln 108
Db 62 TCCAAGCTGCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTGTAGCCGATGCCAA 121
QY 109 GlyLysHisArgArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGlu 128
Db 122 GGAAGCATTAGGAGGTTTGAATGGACCGGAACTAAGAGTCCAGATATGTGTCTGAG 181
QY 129 CysAsnArgGluHisProAlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGly 148
Db 182 TGTAAATAGGCTGCATCTCTGCTGAGGAAGGAGATCTTTTGGCAGAGTCAAGCATGTTGGGC 241
QY 149 LeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAla 168
Db 242 CTCAGATCACCTACTTTGCACTGATGGATGGAAAGGTGTATGACATCACAGAGTGGGCT 301
QY 169 GlyCysGlnArgValGlyLysSerProAspThrHisArgValProTyrHisIleSerPhe 188
Db 302 GGATGCCAGCGTGTAGTATCTCTCCAGATATCCACAGAGTCCCTTATCATCATCTCTTT 361
QY 189 GlySerArgIleProGlyThrArgGlyArgGlnArgAlaThrProAspAlaProAla 208
Db 362 GGTTCCTGGATTCCAGGCCACAGAGGGCGGACAGAGCCACCCAGATGCCCTCTCTGCT 421
QY 209 AspLeuGlnAspPheLeuSerArgIlePheGlnValProGlyGlnMetProAsnGly 228
Db 422 GATCTTCAGATTTCTTTGAGTGGATCTTTCAAGTACCCCGGCGAGATGCCAAT-GGG 480
QY 229 AsnPhePheAlaAlaProGlnProAlaProGlyValAlaAlaAlaSerLysProAsnSer 248
Db 481 AACTCTTTTGCAGCTCTCTCAGCCCTGCCCTGGAGCGCTGCAGCCCTTAGAGCCCAACAGC 540
QY 249 ThrValProLysGlyGluAlaLysProLysArgArgLysLysValArgArgProPheGln 268
Db 541 ACAGTACCACAGGAGAGAGCCAACTTAAGCGCGGAGAAAGAGTGGAGAGGCCCTTCCAA 600
QY 269 Arg 269
Db 601 CGT 603

RESULT 8

AAS30481
ID AAS30481 standard; DNA; 7453 BP.

XX AAS30481;

XX 21-NOV-2001 (first entry)

XX DNA encoding novel prostate gland antigen, Seq ID No 339.

XX Human; neutrotropic; neuroprotective; cytostatic; antiparkinsonian;
KW antianemic; dermatological; immunosuppressive; antineoplastic;
KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;
KW osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;
KW prostatic; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;
KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;
KW reproductive system disorder; autoimmune disorder; urinary system;
KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;
KW blood-related disorder; hyperproliferative disorder; respiratory;
KW neurological disorder; endocrine disorder; inflammatory disorder;
KW liver disorder; wound healing; food preservative; ds.
XX Homo sapiens.
OS

```
XX WO200155447-A1.
PN 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01330.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 10-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234397.
PR 25-SEP-2000; 2000US-0234398.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
XX
XX 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
```


Db 3494 GCGGAAGAAAGTGAGGAGGCCCTTCCAACGT 3524
RESULT 9
AAL06261
ID AAL06261 standard; DNA; 7453 BP.
XX AC
AC AAL06261;
XX XX
DT 21-NOV-2001 (first entry)
XX Human reproductive system related antigen DNA SEQ ID NO: 8949.
DE Human reproductive system related antigen; reproductive system disorder;
KW Human; reproductive system related antigen; cancer; gene therapy; ds.
XX cancer; gene therapy; ds.
XX Homo sapiens.
OS
XX WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180828.
PR 24-FEB-2000; 2000US-0184564.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
PS Disclosure; SEQ ID NO 8949; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX
SQ Sequence 7453 BP; 1980 A; 1633 C; 1692 G; 2147 T; 1 other;

Alignment Scores:
Pred. No.: 2-23e-78 Length: 7453
Score: 1012.50 Matches: 231
Percent Similarity: 45.21% Conservative: 0
Best Local Similarity: 45.21% Mismatches: 2
Query Match: 70.66% Indels: 280
DB: 22 Gaps: 4

US-10-049-742-11 (1-269) x AAL06261 (1-7453)

QY 37 ValHisProAspLysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAla 56
DB 1996 GTTCACTCTGACAAAATCATCATCCCGGCTGAGGAGCCCTTCAAGGTTTTCGGAGCA 2055

QY 57 AlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMet----- 72
DB 2056 GCTTGGGACATTCTCAGCAATGCTGAAAAAGCGAAAGGAGTATGAGAT-GTAAAGTTGGAGA 2114

QY 72 ----- 72
DB 2115 TGGGAAATCATAGATAATGGTAAATGAAATAATCCTCAATAGCAGAGGCATCTGGACTTG 2174

QY 72 ----- 72
DB 2175 GGGGTGGAGGCTTGTGAGATGAGAGAACTGAAGTCACTTGCTTTCTCGCTAGACAGG 2234

QY 72 ----- 72
DB 2235 GGCCCTCAGAGGCCAACTGATATGTCTCTCTTTGTCCCTCCCTCAATACCTTCTGACTTA 2294

QY 73 -----LysArgMetAlaGluAsnGluLeuSerArgSerValAsn 85
DB 2295 CAAAGCATTTCTTCTATTAGGAACCGAAATGGCAGAGAATGAGCTGAGCCGCTCAGTAAAT 2354

QY 86 GluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMetCysSer 105
DB 2355 GAGTTTCTGTCCAAGCTGCAAGATGACCTCAAGAGGCAATGAATACTATGATGTGTAGC 2414

QY 106 ArgCysGlnGlyLysHis----- 111
DB |||||

Db 2415 CGATGCCAAGGAAAGCATAGGTATGAATAGAGGAGAGGGATGGGACAATCAAGCTCA 2474
QY 111 ----- 111
Db 2475 GGGATTATGTAAACCAAGGATCTTTCTAGTAGAGTACTGGTAATTATGAGTTACATCTGTAT 2534
QY 111 ----- 111
Db 2535 CTCATAATTAGTAAAAAGACCCCTTAAGTTTCTACCTTTGTGCATATTGTCTCAATAGACTGAC 2594
QY 111 ----- 111
Db 2595 GATTTTAATTAGCACTAAGAAACATTAGCTTTTAGGATAGAACCTTGAAGGATATGAAAT 2654
QY 111 ----- 111
Db 2655 ACTAAGTGGGTAATTTTAGATGTTTTCAGCCCTATTTTGTGAATAATGGTAAGACTACTAG 2714
QY 112 -----AspArg 113
Db 2715 CAGTGCCTAGGTAGGCTTGACAAATTTAAGGAACATACTCCTTTTGGTTGGTCAGAGAG 2774
QY 113 gPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLeuHi 133
Db 2775 GTTTGAATGGACCGGGAACCTAAGAGTGCCAGATACTGTGCTGAGTGTAAATAGGCTGCA 2834
QY 133 sProAlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGlyLeuLysIleThrTy 153
Db 2835 TCCTGCTGAGGAAGGAGACTTTTGGCAGAGTCAAGCATGTTGGCGCTCAAGATCACCTA 2894
QY 153 rPheAlaLeuMetAspGlyLysValTyrAspIleThr----- 165
Db 2895 CTTTGCACTGATGGATGGAAGGTGTATGACATCACAGGACTCTCTGTCTCTTAGAAAT 2954
QY 165 ----- 165
Db 2955 ACAGGCTCATACTTCTGTATCTTTTATGTGCTCTAAACTCGCAGCTTTTGGCTAATTATT 3014
QY 166 -----GluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisAr 181
Db 3015 TCTCATGTTTATAGATGGCTGGATGCCAGCGTGTAGGTATCTCCCAATATACCCACAG 3074
QY 181 gValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArgAl 201
Db 3075 AGTCCCTCATCATCTCATTTGGTTCTCGGATTCAGGACCAGAGCGGCGGAGAGGTA 3134
QY 201 a----- 201
Db 3135 -GGTGGTATTTCTGTCAATAATCTATCCACTATTTCAGTTTTCAGTTTTCAGATGA 3193
QY 201 ----- 201
Db 3194 CTGCTTTTTCAGACACCCAGGGGCTTGTTCCTAGGAAGTTTGGGAACGTAGTAT 3253
QY 201 ----- 201
Db 3254 ATCTAACTAGGTAAACCATATGACTTAACATCTCTTGCCTTATTTCTTCTGTTTTACCT 3313
QY 202 -----ThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGl 219
Db 3314 CAGAGCCACCCAGATGCCCTCTCTGCTGATCTTCAGGATTTCTTGTAGTCGGATCTTTCA 3373
QY 219 nValProGlyGluMetProAsnGlyAsnPheAlaAlaProGlnProAlaProGl 239
Db 3374 AGTACCCCGAGGCGAGATGCCAATGGGAACCTTTTTCAGCTCTCTCAGCTCCTCAGCGCTG 3433
QY 239 yAlaAlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysAr 259
Db 3434 AGCCGCTGAGCGCTCTTAAGCCCAACAGCAGTACCCAGGGAGAGCCCAACCTAAGCG 3493
QY 259 gArgLysLysValArgArgProPheGlnArg 269
Db 3494 GCGGAAGAAAGTGAGGAGGCCCTTCCACGT 3524

Qy 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122

Alignment Scores:
Pred. No.: 5,698-33 Length: 12578
Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 21 Gaps: 0

US-10-049-742-11 (1-269) x AAZ36196 (1-12578)

QY 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122
DB 5054 GCCAGATACTGCTGAGTGAATAGGCTGCATCTGCTGAGGAAAGGACATTTGGGCA 5113
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
DB 5114 GAGTCGAGCATGTTGGGCTCAAAATCACCTACTTTGGCTGATGATGGAAGGTGTAT 5173
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
DB 5174 GATATCACAGAGTGGGCTGAGTGGCAGCGTGTGGGAATCTCCACAGATACCCACAGATC 5233
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
DB 5234 CCTGTGCATCTCATTTGGTTACGGATGCT 5266

RESULT 12
ABA95615
ID ABA95615 standard; DNA; 12734 BP.
XX ABA95615;
AC ABA95615;
DT 21-MAR-2002 (first entry)
XX
DE Chimeric BVDV/HCV NS3-wt sequence.
XX
KW Pestivirus; Npro; protease; NS3; screening; ds.
XX
OS Chimeric - Bovine viral diarrhea virus.
OS Chimeric - Hepatitis C virus.
XX US6326137-B1.
XX
PD 04-DEC-2001.
XX
PF 25-JUN-1999; 99US-0344456.
XX
PR 25-JUN-1999; 99US-0344456.
XX (SCHE) SCHERING CORP.
XX
PI Hong Z, Lai VCH, Lau JYN;
XX
DR WPI; 2002-121103/16.
XX
PT Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)
PT pestivirus genome where the Npro protease gene is replaced with NS3
PT protease gene, useful for in vivo screening of compounds which inhibit
PT HCV infection -
XX
PS Example 2; Columns 17-28; 20pp; English.
XX
CC The present invention relates to a nucleic acid construct encoding a
CC chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct
CC comprises a pestivirus genome where a Npro pestivirus protease gene is
CC replaced with a gene encoding a functional HCV NS3 protease. Furthermore,
CC each junction site recognised by the Npro protease is replaced with a
CC junction site recognised by the HCV NS3 protease. The construct is useful

CC for screening compounds that inhibit HCV in vivo by inhibiting HCV
CC protease, where screening may be in cell culture or in an animal model.
CC The present sequence is a chimeric clone of BVDV (bovine viral diarrhea
CC virus)/HCV NS3-wt, which was used to illustrate the present invention.
XX
SQ Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;
Alignment Scores:
Pred. No.: 5,788-33 Length: 12734
Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 24 Gaps: 0

US-10-049-742-11 (1-269) x ABA95615 (1-12734)

QY 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122
DB 5150 ATGTCCACCCGATGCCAGGAAAGCATAGGAGTTTGAATGACCGGAACTTAAGAGT 5209
QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAla 142
DB 5210 GCCAGATACTGCTGAGTGAATAGGCTGCATCTGCTGAGGAAAGGTGACITTTGGGCA 5269
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
DB 5270 GAGTCGAGCATGTTGGGCTCAAAATCACCTACTTTGGCTGATGATGGAAGGTGTAT 5329
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
DB 5330 GATATCACAGAGTGGGCTGAGTGGCAGCGTGTGGGAATCTCCACAGATACCCACAGATC 5389
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
DB 5390 CCTGTGCATCTCATTTGGTTACGGATGCT 5422

RESULT 13
AAZ36203
ID AAZ36203 standard; DNA; 12842 BP.
XX AAZ36203;
AC AAZ36203;
DT 11-FEB-2000 (first entry)
XX
DE Nucleotide sequence of the prototype HCV-BVDV chimera.
XX
KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;
KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;
KW bovine viral diarrhea virus; NADL; vaccine; ss.
XX Chimeric - Hepatitis C virus.
OS Chimeric - Bovine viral diarrhea virus.
XX WO9955366-A1.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-US08850.
XX
PR 24-APR-1998; 98US-0082964.
XX (UNIW) UNIV WASHINGTON.
XX
PI Rice CM, Frolov I, McBride MS;
XX WPI; 2000-013359/01.
XX
PT Chimeric viral RNA, used in vaccine against BVDV -
XX Example 5; Fig 19; 108pp; English.
XX
CC The present sequence represents the prototype Hepatitis C virus

(HCV)-bovine viral diarrhea virus (BVDV) chimeric virus of the invention. The sequence contains the adapted HCV 5'NTR from 5'NTR/R.3orig and tandem 3'NTR elements from HCV followed by BVDV. The specification describes chimeric viral RNA comprising a 5' nontranslated region (5'NTR); an open reading frame (ORF) region; and a 3' NTR; where at least one of the regions is chimeric and comprises a nucleotide sequence from a pestivirus in operable linkage with a heterologous nucleotide sequence, preferably from HCV. The chimeric viral RNA is replication-competent. The chimeric viral RNA can be used in a method for identifying compounds having antiviral activity against HCV. When the pestivirus viral nucleotide sequence is from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be used in a vaccine against BVDV.

XX Sequence 12842 BP; 4034 A; 2612 C; 3282 G; 2914 T; 0 other;

Alignment Scores:
Pred. No.: 5,84e-33 Length: 12842
Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 21 Gaps: 0

US-10-049-742-11 (1-269) x AAZ36203 (1-12842)

QY 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122
Db 4962 ATGTGAGCGGATGCCAGGAAAGCATAGAGGTTTGAATAGCAGCGGAACCTAAGAGT 5021
QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPheTrpAla 142
Db 5022 GCCAGATACGTGCTGAGTGAATAGCTGCATCTCTGAGGAAAGGTGACATTTGGGCA 5081
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
Db 5082 GAGTCGAGCATGTGGGCTCTCAAAATCACCCTACTTTGCCGCTGATGGATGGAAGGTGTAT 5141
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
Db 5142 GATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCACAGAGTC 5201
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
Db 5202 CCTTGTGCATCTCATTTGGTTTCACGGATGCCT 5234

RESULT 14
AAZ36211
ID AAZ36211 standard; DNA; 13198 BP.
XX
AC AAZ36211;
XX
DT 11-FEB-2000 (first entry)
DE Nucleotide sequence of functional HCV-BVDV chimera from pCBV/p7/IRES-pac.
XX
KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;
KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;
KW bovine viral diarrhea virus; NADL; vaccine; ss.
XX
OS Chimeric - Hepatitis C virus.
OS Chimeric - Bovine viral diarrhea virus.
XX
PN WO995366-A1.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-US08850.
XX
PR 24-APR-1998; 98US-0082964.
XX
PA (UNIW) UNIV WASHINGTON.
XX

PI Rice CM, Frolov I, McBride MG;
XX
DR WPI; 2000-013359/01.
XX
PT Chimeric viral RNA, used in vaccine against BVDV -
XX
PS Example 3; Fig 24; 108pp; English.
CC The present sequence represents a functional Hepatitis C virus (HCV)-bovine viral diarrhea virus (BVDV) chimeric virus of the invention, expressing a dominant selectable marker conferring resistance to puromycin. The specification describes chimeric viral RNA comprising a 5' nontranslated region (5'NTR); an open reading frame (ORF) region; and a 3' NTR; where at least one of the regions is chimeric and comprises a nucleotide sequence from a pestivirus in operable linkage with a heterologous nucleotide sequence, preferably from HCV. The chimeric viral RNA is replication-competent. The chimeric viral RNA can be used in a method for identifying compounds having antiviral activity against HCV. When the pestivirus viral nucleotide sequence is from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be used in a vaccine against BVDV.

XX Sequence 13198 BP; 3735 A; 3097 C; 3533 G; 2833 T; 0 other;

Alignment Scores:
Pred. No.: 6,04e-33 Length: 13198
Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 21 Gaps: 0

US-10-049-742-11 (1-269) x AAZ36211 (1-13198)

QY 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122
Db 4089 ATGTGAGCGGATGCCAGGAAAGCATAGAGGTTTGAATAGCAGCGGAACCTAAGAGT 4148
QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPheTrpAla 142
Db 4149 GCCAGATACGTGCTGAGTGAATAGCTGCATCTCTGAGGAAAGGTGACATTTGGGCA 4208
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
Db 4209 GAGTCGAGCATGTGGGCTCAAAATCACCCTACTTTGCCGCTGATGGATGGAAGGTGTAT 4268
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
Db 4269 GATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCACAGAGTC 4328
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
Db 4329 CCTTGTGCATCTCATTTGGTTTCACGGATGCCT 4361

RESULT 15
AAZ38807
ID AAZ38807 standard; DNA; 14078 BP.
XX
AC AAZ38807;
XX
DT 31-AUG-2000 (first entry)
DE Plasmid pBVdDn1, carrying bovine viral diarrhea virus, BVdDn1.
XX
KW Bovine viral diarrhea virus; attenuation; Npro protease; enzyme;
KW mucosal disease; ulcer; enteric disease; pneumonia; BVdDn1; antiviral;
KW diarrhoea; ds.
XX
OS Bovine viral diarrhea virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 39..12116

Job time : 359 secs

FT /tag= a
FT /label= BVDvN1_genome
FT /note= "BVDvN1 genome is claimed in Claim 1 of
FT specification"

XX EP1013757-A2.
PN
XX
XX 28-JUN-2000.
PD
XX
XX 08-NOV-1999; 99EP-0308866.
PF
XX
XX 10-NOV-1998; 98US-0107908.
PR
XX
XX (PFIZ) PFIZER PROD INC.
PA
XX
XX Cao X, Sheppard MG;
PI
XX
XX WPI; 2000-414600/36.
DR
XX
XX Attenuated bovine viral diarrhoea virus, used as a vaccine to give
PT cattle protective immunity against subsequent infection with the virus
PT .
XX
XX Claim 1; Fig 2; 44pp; English.
PS
XX
XX Bovine viral diarrhoea virus (BVDV) causes mucosal disease in infected
CC cattle. Symptoms include elevated temperature, coughing, diarrhoea and
CC ulceration of the alimentary mucosa. This virus can be transmitted via
CC the placenta to unborn calves. These calves suffer from persistent
CC infection and are highly predisposed to infection with microorganisms
CC causing diseases such as pneumonia or enteric diseases. The present
CC sequence is a plasmid which carries a mutated form of the bovine viral
CC diarrhoea virus (BVDVdN1). The mutated BVDV was produced by deleting the
CC Npro protease gene from the wild type viral genome. The Npro gene
CC functions in proteolytic cleavage. The mutated virus was found to be
CC attenuated i.e. the virus replicates at a slower rate than the wild type
CC virus and so is less infectious. The mutated virus is therefore suitable
CC for use in vaccines for cattle.
XX
SQ Sequence 14078 BP; 4407 A; 2901 C; 3608 G; 3162 T; 0 other;

Alignment Scores:
Pred. No.: 6.55e-33 Length: 14078
Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 21 Gaps: 0

US-10-049-742-11 (1-269) x AAA38807 (1-14078)

QY 103 MetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluProLysSer 122
Db |||||
4532 ATGTGCAGCCGATGCCAGGAAAGCATAGGAGGTTGAAATGGACCGGAACTAAGAGT 4591

QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPheTpaAla 142
Db |||||
4592 GCCAGATACTGTGCTGAGTGTAAATAGGCTGCATCTCTGCGAGGAGGTGACTTTGGGCA 4651

QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
Db |||||
4652 GAGTCGAGCATGTTGGGCCCTCAAAATCACCTACTTTGCCCTGATGGATGGAAGGTGTAT 4711

QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
Db |||||
4712 GATATCACAGATGGGCTGGATGCCAGCGTGTGGGAATCTCCCGAGATACCCACAGAGTC 4771

QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
Db |||||
4772 CCTTGTCACATCTCATTTGGTTTCACGGATGCCT 4804

Search completed: December 1, 2003, 10:16:02

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 1, 2003, 09:37:08 ; Search time 74 Seconds

(without alignments)
1604.487 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433

Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPGKAPKRRKRVRRPQR 269

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Maximum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10049742/runat_01122003_073137_18627/app_query.fasta_1.455
-DB=Issued Patents NA -QEXT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10049742@cgn_1_1_56@runat_01122003_073137_18627 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	497	34.7	12734	4	US-09-344-456-1		Sequence 1, Appli
2	497	34.7	14078	3	US-09-433-262-1		Sequence 1, Appli
3	497	34.7	14078	4	US-09-702-330-1		Sequence 1, Appli
4	497	34.7	14578	3	US-08-859-694-1		Sequence 1, Appli
5	173.5	12.1	1128	4	US-09-328-352-858		Sequence 858, App
6	159	11.1	635	4	US-09-669-751-166		Sequence 166, App
7	156.5	10.9	2349	2	US-08-974-546-2		Sequence 2, Appli
8	155.5	10.9	1376	2	US-08-868-288A-2		Sequence 2, Appli
9	155.5	10.9	1376	3	US-09-235-373-2		Sequence 2, Appli
10	155.5	10.9	1376	3	US-09-388-993-2		Sequence 2, Appli
11	155.5	10.9	1621	4	US-09-996-243-147		Sequence 147, App
12	155	10.8	8050	3	US-09-491-362-11		Sequence 11, Appli

13	155	10.8	8050	4	US-09-874-562-11		Sequence 11, Appli
14	152.5	10.6	4403765	3	US-09-103-840A-2		Sequence 2, Appli
15	152.5	10.6	4411529	3	US-09-103-840A-1		Sequence 1, Appli
16	151.5	10.6	1146	4	US-09-252-991A-10603		Sequence 10603, A
17	151.5	10.6	1260	4	US-09-252-991A-10250		Sequence 10250, A
18	149.5	10.4	1014	4	US-09-658-644-3		Sequence 3, Appli
19	149.5	10.4	5687	4	US-09-221-017B-368		Sequence 368, App
20	146	10.2	1516	4	US-09-149-476-188		Sequence 188, App
21	144	10.0	360	4	US-09-370-838-140		Sequence 140, App
22	144	10.0	503	4	US-09-370-838-165		Sequence 165, App
23	143.5	10.0	1482	4	US-09-613-303-18		Sequence 18, Appli
24	143.5	10.0	403765	3	US-09-103-840A-2		Sequence 2, Appli
25	143.5	10.0	4411529	3	US-09-103-840A-1		Sequence 1, Appli
26	143	10.0	42931	4	US-08-311-731A-129		Sequence 129, App
27	143	10.0	580073	4	US-08-545-528D-1		Sequence 1, Appli
28	141.5	9.9	1230025	4	US-09-198-452A-1		Sequence 1, Appli
29	140	9.8	1881	4	US-09-553-498-1		Sequence 1, Appli
30	140	9.8	1881	4	US-09-553-498-3		Sequence 3, Appli
31	140	9.8	1881	4	US-09-618-869-1		Sequence 1, Appli
32	140	9.8	1881	4	US-09-618-869-3		Sequence 3, Appli
33	137	9.6	963	4	US-09-328-352-3811		Sequence 3811, Ap
34	134	9.4	1330	2	US-08-868-288A-4		Sequence 4, Appli
35	134	9.4	1330	3	US-09-235-373-4		Sequence 4, Appli
36	134	9.4	1330	3	US-09-388-993-4		Sequence 4, Appli
37	133	9.3	1158	4	US-09-134-001C-851		Sequence 851, App
38	131.5	9.2	672	1	US-08-486-955A-6		Sequence 6, Appli
39	130.5	9.1	1830121	4	US-09-557-884-1		Sequence 1, Appli
40	130.5	9.1	1830121	4	US-09-643-990A-1		Sequence 1, Appli
41	129.5	9.0	1767	4	US-09-620-312D-820		Sequence 820, App
42	128.5	9.0	509	4	US-09-996-243-149		Sequence 149, App
43	128.5	9.0	1119	4	US-09-252-991A-10788		Sequence 10788, A
44	126.5	8.8	4320	2	US-08-472-534-4		Sequence 4, Appli
45	126.5	8.8	7174	4	US-08-961-527-189		Sequence 189, App

ALIGNMENTS

RESULT 1
US-09-344-456-1
; Sequence 1, Application US/09344456A
; Patent No. 6326137
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Lai, Vicki C.H.
; APPLICANT: Lau, Johnson Y.N.
; TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE-DEPENDENT CHIMERIC
; TITLE OF INVENTION: PESTIVIRUS
; FILE REFERENCE: IN01038
; CURRENT APPLICATION NUMBER: US/09/344,456A
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 12734
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: Pestivirus
US-09-344-456-1

Alignment Scores:
Pred. No.: 1,86e-37 Length: 12734
Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 4 Gaps: 0

US-10-049-742-11 (1-269) x US-09-344-456-1 (1-12734)

QY 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122
|||||

Db 5150 ATGTGACCGCATGCCAGGAAAGCATAGGAGTTTGAATGACCGGAACTAAGAGT 5209
QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAla 142
Db 5210 GCCAGATACTGCTGAGTGTAATAGCTGCATCTCTGAGGAAGGTGACTTTTGGGCA 5269
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
Db 5270 GAGTCGACGATGTTGGCGCTCAAAATCACCTACTTTCGCTGATGGAAGGTGTAT 5329
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
Db 5330 GATATCACAGAGTGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCACAGAGTC 5389
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
Db 5390 CTTGTGCATCTCATTTGGTTTACGGATGCCT 5422

RESULT 2

US-09-433-262-1
; Sequence 1, Application US/09433262
; Patent No. 6168942
; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; TITLE OF INVENTION: ATTENUATED FORMS OF BOVINE VIRAL DIARRHEA VIRUS
; FILE REFERENCE: PC10435A
; CURRENT APPLICATION NUMBER: US/09/433,262
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 60/107,908
; EARLIER FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 14078
; TYPE: DNA
; ORGANISM: Bovine Viral Diarrhea Virus
US-09-433-262-1

Alignment Scores:
Pred. No.: 2,14e-37 Length: 14078

Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 3 Gaps: 0

US-10-049-742-11 (1-269) x US-09-433-262-1 (1-14078)

QY 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122
Db 4532 ATGTGACCGCATGCCAGGAAAGCATAGGAGTTTGAATGACCGGAACTAAGAGT 4591
QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAla 142
Db 4592 GCCAGATACTGCTGAGTGTAATAGCTGCATCTCTGAGGAAGGTGACTTTTGGGCA 4651
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
Db 4652 GAGTCGACGATGTTGGCGCTCAAAATCACCTACTTTCGCTGATGGAAGGTGTAT 4711
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
Db 4712 GATATCACAGAGTGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCACAGAGTC 4771
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
Db 4772 CTTGTGCATCTCATTTGGTTTACGGATGCCT 4804

RESULT 3

US-09-702-330-1
; Sequence 1, Application US/09702330
; Patent No. 6410032

; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; TITLE OF INVENTION: ATTENUATED FORMS OF BOVINE VIRAL DIARRHEA VIRUS
; FILE REFERENCE: PC10435A
; CURRENT APPLICATION NUMBER: US/09/702,330
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/433,262
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 14078
; TYPE: DNA
; ORGANISM: Bovine Viral Diarrhea Virus
US-09-702-330-1

Alignment Scores:
Pred. No.: 2,14e-37 Length: 14078

Score: 497.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 4 Gaps: 0

US-10-049-742-11 (1-269) x US-09-702-330-1 (1-14078)

QY 103 MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer 122
Db 4532 ATGTGACCGCATGCCAGGAAAGCATAGGAGTTTGAATGACCGGAACTAAGAGT 4591
QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAla 142
Db 4592 GCCAGATACTGCTGAGTGTAATAGCTGCATCTCTGAGGAAGGTGACTTTTGGGCA 4651
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
Db 4652 GAGTCGACGATGTTGGCGCTCAAAATCACCTACTTTCGCTGATGGAAGGTGTAT 4711
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
Db 4712 GATATCACAGAGTGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCACAGAGTC 4771
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
Db 4772 CTTGTGCATCTCATTTGGTTTACGGATGCCT 4804

RESULT 4

US-08-859-694-1
; Sequence 1, Application US/08859694A
; Patent No. 6001613
; GENERAL INFORMATION:
; APPLICANT: Donis, Ruben O.
; APPLICANT: Vassilev, Ventsislav B.
; TITLE OF INVENTION: A plasmid bearing a cDNA copy of the genome of bovine
; TITLE OF INVENTION: viral diarrhea virus, chimeric derivations thereof, and
; TITLE OF INVENTION: method of producing an infectious bovine viral diarrhea
; TITLE OF INVENTION: virus using said plasmid
; FILE REFERENCE: UNVNS1110
; CURRENT APPLICATION NUMBER: US/08/859,694A
; CURRENT FILING DATE: 1997-05-21
; EARLIER APPLICATION NUMBER: 60/018,246
; EARLIER FILING DATE: 1996-05-24
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 14578
; TYPE: DNA
; ORGANISM: bovine viral diarrhea virus
US-08-859-694-1

Alignment Scores:
Pred. No.: 2,24e-37 Length: 14578

```

Score: 457.00 Matches: 89
Percent Similarity: 98.90% Conservative: 1
Best Local Similarity: 97.80% Mismatches: 1
Query Match: 34.68% Indels: 0
DB: 3 Gaps: 0

US-10-049-742-11 (1-269) x US-08-859-694-1 (1-14578)

QY 103 MetCysSerArgCysGlnGlyLysHisAArgPheGluMetAspArgGluProLysSer 122
Ddb 5032 ATGTGCAGCGCATGCCAGGAAAGCATAGCAGGTTTGAATGCACCGGAAACCTAAGAGT 5091
QY 123 AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAla 142
Ddb 5092 GCCAGATACTGTGCTCAGTGTGAATAGCGTCATCCTGCTGAGGAAGTCACTTTTGGGCA 5151
QY 143 GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr 162
Ddb 5152 GAGTGCAGCATGTGGCCCTCAAAATCACTACTTTTGGCTGATGGATGGAAGGTGTAT 5211
QY 163 AspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal 182
Ddb 5212 GATATCACAGAGTGGCTCGATGCCAGCGTGTGGGAATCTCCACAGATATCCACACAGTC 5271
QY 183 ProTyrHisIleSerPheGlySerArgIlePro 193
Ddb 5272 CCTTGTGCATCTCATTTGGTTTCACGGATGCCT 5304

RESULT 5
US-09-328-352-858
; Sequence 858, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 858
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (965)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-328-352-858

Alignment Scores:
Pred. No.: 9,67e-08 Length: 1128
Score: 173.50 Matches: 63
Percent Similarity: 36.02% Conservative: 31
Best Local Similarity: 24.14% Mismatches: 84
Query Match: 12.11% Indels: 83
DB: 4 Gaps: 11

US-10-049-742-11 (1-269) x US-09-328-352-858 (1-1128)

QY 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
Ddb 34 TATGAGTTTTTAGCGTTTCAAAAACCGCAAGTGTATGATGATCAAAAAAGCCTATCGT 93
QY 32 GlnLeuAlaValMetValHisProAspLysAsn---HisHisProArgAlaGluAla 50
Ddb 94 AAATTTGGCGATGAAATATATCATCTCTCACAGAAAACCCCTGACAATGCCGAGGCTGAAGAAAAA 153
QY 51 PheLysValLeuArgAlaAlaThrAspIleValSerAsnAlaGluLysArgLysGluTyr 70
Ddb 154 TTTAAAGAAAGCTTCTGAAGCTTATGAGATTTTATCGGACAGCGAAAAACGCAGCATGTAT 213
QY 71 GluMetLysArgMetAlaGluAsnGluLeuSerArg----- 82

```

```
US-10-049-742-11 (1-269) x US-09-669-751-166 (1-635)
QY 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
DB 186 TACAAGATTCTGGGCTCGAGCGCAAGCGCGACGATGAGATCAAGAAGCGCTACCGC 245
QY 32 GlnLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPhe 51
DB 246 AAATCTGGCACTCAATACATCCCGACAGAACAGAGCCACAGCGGAGAGCGCTTC 305
QY 52 LysValLeuArgAlaAlaTirAspIleValSerAsnAlaGluLysArgLysGluTyrGlu 71
DB 306 AAGGAGATCGCGAGCGGTACGAGGTGTGTGCGGACAAAAGAGCGGACATCTTCGAC 365

RESULT 7
US-08-974-546-2
; Sequence 2, Application US/08974546
; Patent No. 5945287
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: TWO HUMAN HEAT SHOCK PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,546
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2349 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT21
; CLONE: 2525691
; US-08-974-546-2

Alignment Scores:
Pred. No.: 1,15e-05 Length: 2349
Score: 156.50 Matches: 66
Percent Similarity: 38.61% Conservative: 34
Best Local Similarity: 25.48% Mismatches: 79
Query Match: 10.92% Indels: 80
DB: 2 Gaps: 9

US-10-049-742-11 (1-269) x US-08-974-546-2 (1-2349)
QY 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
DB 186 TACAAGATTCTGGGCTCGAGCGCAAGCGCGACGATGAGATCAAGAAGCGCTACCGC 245
QY 32 GlnLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPhe 51
DB 246 AAATCTGGCACTCAATACATCCCGACAGAACAGAGCCACAGCGGAGAGCGCTTC 305
QY 52 LysValLeuArgAlaAlaTirAspIleValSerAsnAlaGluLysArgLysGluTyrGlu 71
DB 306 AAGGAGATCGCGAGCGGTACGAGGTGTGTGCGGACAAAAGAGCGGACATCTTCGAC 365

RESULT 8
US-08-868-288A-2
; Sequence 2, Application US/08868288A
; Patent No. 5922567
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: TWO NEW HUMAN DNAJ-LIKE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2349 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT21
; CLONE: 2525691
; US-08-974-546-2
```

```

; APPLICATION NUMBER: US/08/868,288A
; FILING DATE: June 3, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0309 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAB01
; CLONE: 136466
US-08-868-288A-2

Alignment Scores:
Pred. No.: 6,89e-06 Length: 1376
Score: 155.50 Matches: 29
Percent Similarity: 80.33% Conservative: 20
Best Local Similarity: 47.54% Mismatches: 11
Query Match: 10.85% Indels: 1
DB: 2 Gaps: 1

US-10-049-742-11 (1-269) x US-08-868-288A-2 (1-1376)
QY 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
DB 274 TATAAGATCTTGGGGTGGCTCGAGTGCCTCTATTAAGGATATTAAAGGCTATAGG 333
QY 32 GlnLeuAlaValMetValHisProAspLysAsn---HisHisProArgAlaGluGluAla 50
DB 334 AAACCTAGCCCTCCAGCTTCATCCGACCGGACCCCTGATGATCCACAGCCGAGGAAA 393
QY 51 PheLysValLeuArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyr 70
DB 394 TTCCAGGATCTGGGTGCTCTATGAGGTTCCTCAGATAGTGCAGAAACGGAACAGTAC 453
QY 71 Glu 71
DB 454 GAT 456

RESULT 9
US-09-235-373-2
; Sequence 2, Application US/09235373
; Patent No. 6001598
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: TWO NEW HUMAN DNAJ-LIKE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,373

```

```

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/868,288
; FILING DATE: June 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0309 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAB01
; CLONE: 136466
US-09-235-373-2

Alignment Scores:
Pred. No.: 6,89e-06 Length: 1376
Score: 155.50 Matches: 29
Percent Similarity: 80.33% Conservative: 20
Best Local Similarity: 47.54% Mismatches: 11
Query Match: 10.85% Indels: 1
DB: 3 Gaps: 1

US-10-049-742-11 (1-269) x US-09-235-373-2 (1-1376)
QY 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
DB 274 TATAAGATCTTGGGGTGGCTCGAGTGCCTCTATTAAGGATATTAAAGGCTATAGG 333
QY 32 GlnLeuAlaValMetValHisProAspLysAsn---HisHisProArgAlaGluGluAla 50
DB 334 AAACCTAGCCCTCCAGCTTCATCCGACCGGACCCCTGATGATCCACAGCCGAGGAAA 393
QY 51 PheLysValLeuArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyr 70
DB 394 TTCCAGGATCTGGGTGCTCTATGAGGTTCCTCAGATAGTGCAGAAACGGAACAGTAC 453
QY 71 Glu 71
DB 454 GAT 456

RESULT 10
US-09-388-993-2
; Sequence 2, Application US/09388993
; Patent No. 6043222
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: TWO NEW HUMAN DNAJ-LIKE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/388,993
; FILING DATE:

```

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/868,288
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0309 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1376 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: SYNORAB01
/ CLONE: 136466
/ US-09-388-993-2

Alignment Scores:
Pred. No.: 6,89e-06 Length: 1376
Score: 155.50 Matches: 29
Percent Similarity: 80.33% Conservative: 20
Best Local Similarity: 47.54% Mismatches: 11
Query Match: 10.85% Indels: 1
DB: 3 Gaps: 1

US-10-049-742-11 (1-269) x US-09-388-993-2 (1-1376)
QY 12 PheHisValIeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
Db 274 TATAAGATCTGGGGTGGCTCGAAGTGCCTCTATAAAGGATATTAAGAGGCTATAGG 333
QY 32 GinIeuAlaValMetValHisProAspLysAsn--HisHisProArgAlaGluGluAla 50
Db 334 AAACAGCCCTGCAGCTGATCCGACCGGAAACCTGTATGATCCAAAGCCGAGGAAA 393
QY 51 PheLysValIeuArgAlaAlaIleValSerAsnAlaGluLysArgLysGluTyr 70
Db 394 TTCCAGATCTGGGTGCTGCTTATAGGTTCTCTCAGATAGTAGAGAAACGGAACAGTAC 453
QY 71 Glu 71
Db 454 GAT 456

RESULT 11
US-09-996-243-147
/ Sequence 147, Application US/09996243
/ Patent No. 6478825
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi J.
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Kijavini, Ivar J.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
```

```
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ TITLE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCE: P2730P1C13
/ CURRENT APPLICATION NUMBER: US/09/996,243
/ CURRENT FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: 60/049787
/ PRIOR FILING DATE: 1997-06-16
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/065186
/ PRIOR FILING DATE: 1997-11-12
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066770
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/075945
/ PRIOR FILING DATE: 1998-02-25
/ PRIOR APPLICATION NUMBER: 60/078910
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/084600
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/087106
/ PRIOR FILING DATE: 1998-05-28
/ PRIOR APPLICATION NUMBER: 60/087607
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087609
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087759
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/087827
/ PRIOR FILING DATE: 1998-06-03
/ PRIOR APPLICATION NUMBER: 60/088021
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088025
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088026
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088028
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088029
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088030
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088033
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088326
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/089167
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088202
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088212
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088217
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088655
/ PRIOR FILING DATE: 1998-06-09
/ PRIOR APPLICATION NUMBER: 60/088734
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088738
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088742
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088810
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088824
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088826
```

; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 8,63e-06 Length: 1621
Score: 155.50 Matches: 29
Percent Similarity: 80.33% Conservative: 20
Best Local Similarity: 47.54% Mismatches: 11
Query Match: 10.85% Indels: 1
DB: 4 Gaps: 1

US-10-049-742-11 (1-269) x US-09-996-243-147 (1-1621)

Qy 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysAlaTyrArg 31
::: :::::::::::::::::::::
Db 237 TATAAGATCTTGGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGCCCTATAGG 296
::: :::::::::::::::::::::

Qy 32 GlnLeuAlaValMetValHisProAspLysAsn---HisHisProArgAlaGluGluAla 50
::: :::::::::::::::::::::
Db 297 AAACCTAGCCCTGCAGCTTCATCCCGACCGCAACCTGATGATCCACAGCCCGAGAGAAA 356
::: :::::::::::::::::::::

Qy 51 PheLysValLeuArgAlaAlaIleAspIleValSerAsnAlaGluLysArgLysGluTyr 70
::: :::::::::::::::::::::
Db 357 TTCAGGATCTGGGTGCTGCTTATGAGGTTCTGTCAGATAGTGAGAAACGGAACAGTAC 416
::: :::::::::::::::::::::

Qy 71 Glu 71
:::
Db 417 GAT 419

RESULT 12
US-09-491-362-11
; Sequence 11, Application US/09491362
; Patent No. 6281017
; GENERAL INFORMATION:
; APPLICANT: Croceau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: 1-DEOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: WSUR14977
; CURRENT APPLICATION NUMBER: US/09/491,362
; CURRENT FILING DATE: 2000-01-26
; EARLIER APPLICATION NUMBER: 60/118,349
; EARLIER FILING DATE: 1999-02-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0

; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25

```
; SEQ ID NO 11
; LENGTH: 8050
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-491-362-11

Alignment Scores:
Pred. No.:      8,766-05      Length:      8050
Score:          155.00      Matches:      61
Percent Similarity: 38.61%      Conservative: 39
Best Local Similarity: 23.55%      Mismatches: 92
Query Match:      10.82%      Indels:      68
DB:              3          Gaps:      9

US-10-049-742-11 (1-269) x US-09-491-362-11 (1-8050)
QY 6 GluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThrAlaSerAspValGlu 25
Db 1061 GAAGGAGAAATCTGACTGATGGAATCCTTGGTGTGATCCTTTAGCTGATGAAACA 1120
QY 26 LeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAspLysAsnHisPro 45
Db 1121 GTGAAGAAACATTACAAGACCTTAGCTCTGTGCTTCCCGGCAAGAACAGGTTTAAT 1180
QY 46 ArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIleValSerAsnAlaGlu 65
Db 1181 GGTGCGGAAGTGCCTTTAAGCTGCTTTAGATGCTTGTCTCTACTATCTGATAAGCT 1240
QY 66 LysArgLysGluTyrGluMetLysArg-----MetAlaGluAsnGluLeuSerArg 82
Db 1241 AAGAGAATTGC-GTTGATCAAAAGAGAAACCAAAACAAAGAAAGAGCAACCATCTGCT 1299
QY 83 SerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn----- 100
Db 1300 TCGTGTAAATAGCCTGCAGAGCCTCTTCTTCTTCTTCGTCGAAACCGGTGCACATGACC 1359
QY 101 -----ThrMetCysSerArgCysGlnGlyLysHisArg 112
Db 1360 TTTTCGACAGTGACATGACCTTTTCGACAGATGCAATAAATGCACAACGAGATGTTGT 1419
QY 113 ArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLeu 132
Db 1420 CATTTTTCGACGACAGATCATCTTAAACAGACCTTTCCTTGTCCAAACTGTGTCAG--- 1476
QY 133 HisProAlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGlyLeuLysIleThr 152
Db 1477 -----AATTTCGGCTATGACCAATATATCATCGACA 1506
QY 153 TyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAlaGlyCysGlnArg 172
Db 1507 -----GAGGTGATCAATGGGAGGACATTC-----ATCAGA 1536
QY 173 ValGlyIleSerProAspThrHisArgValProTyrHisIleSerPheGlySerArgIle 192
Db 1537 GTCCTCTGTTCTCCGCAACAAAGAGAA-----CCATCGAGGGCC 1575
QY 193 ProGlyThrArgGlyArgGlnAlaThr-----ProAspAlaProProAlaSerLeu 210
Db 1576 AATTCTCAAGCAACTAGCAGACGTAGCAGCATGATGATGCAAACTCTACTGAG--- 1632
QY 211 GlnAspPheLeuSerArgIlePheGlnValProProGlyGlnMetProAsnGlyAsnPhe 230
Db 1633 -----AGTTT 1638
QY 231 PheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerLysProAsnSerThr 249
Db 1639 TTCAGAAACCAATGCCACACAGGAGATGCA-----AATCTTACT 1680

RESULT 13
US-09-874-562-11
; Sequence 11, Application US/09874562
; Patent No. 6420159
; GENERAL INFORMATION:
```

Qy	211	GlnAspPheLeuSerArgIlePheGlnValProProGlyGlnMetProAsnGlyAsn	Phe 230
		:::	
Db	1633	-----AGTTT	1638
Qy	231	PheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerLysProAsnSerThr	249
Db	1639	TTTCAAGAAACCAATCCGACACAGGAGATGCA-----AACTCTACT	1680
RESULT 14			
US-09-103-840A-2/c			
; Sequence 2, Application US/09103840A			
; Patent No. 6294328			
; GENERAL INFORMATION:			
; APPLICANT: FLEISCHMAN, Robert D.			
; APPLICANT: WHITE, Owen R.			
; APPLICANT: FRASER, Claire M.			
; APPLICANT: VENTER, John C.			
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM			
; TITLE OF INVENTION: TUBERCULOSIS			
; FILE REFERENCE: 24366-20007.00			
; CURRENT APPLICATION NUMBER: US/09/103,840A			
; CURRENT FILING DATE: 1998-06-24			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 2			
; LENGTH: 4403765			
; TYPE: DNA			
; ORGANISM: Mycobacterium tuberculosis			
; FEATURE:			
; OTHER INFORMATION: CDC 1551			
; OTHER INFORMATION: "n" bases at various positions throughout the sequence			
; OTHER INFORMATION: represent a, t, c or g			
US-09-103-840A-2			
Alignment Scores:			
Pred. No.:		0.897	Length: 4403765
Score:		152.50	Matches: 79
Percent Similarity:		30.54%	Conservative: 45
Best Local Similarity:		19.46%	Mismatches: 100
Query Match:		10.64%	Indels: 183
DB:		3	Gaps: 14
US-10-049-742-11 (1-269) x US-09-103-840A-2 (1-4403765)			
Qy	12	PreHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTrArg	31
		::: ::	
Db	2651266	TACGGCTGCTCGCGGTGAGCAAGACGGCAGCGACCGGACATCAACCGCCTACCGC	2651207
Qy	32	GlnLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPhe	51
		:::	
Db	2651206	AACTGGCGCGCAGCTGATCCCGACGTCAACCCGGAGCAGGCTCGCAGCGGGAATTC	2651147
Qy	52	LysValLeuArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTrpGlu	71
		::: ::	
Db	2651146	AAGAAATCAGCTCGCTACGAGGTGCTCAGTGACCCGGACAAACGTCGCATCGTCGAC	2651087
Qy	72	MetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPhe	87
		::: ::	
Db	2651086	CTGGCGGGGATCCGCTGGAGAGCGCGCTCGCGGGCAATGGGTTCGGTTCGGTTCGGC	2651027
Qy	87	-----	87
Db	2651026	GGCCTCGGCACGTGTTTCGAGGCGTTCTTTGGCGGGGTTTCGGTGGGGCGCGCGTCC	2650967
Qy	88	-----LeuSerLysLeuGlnAspArg	94
Db	2650966	CGAGGGCGGATCGCGCGGTTCGCGCGGTTCGGATCTCGTCTGCTACGAATGCGGTGGAT	2650907
Qy	95	LeuLysGluAlaMet-----AsnThrMetMetCys	104
Db	2650906	CTCGAAGATGCCCAACAGGTGTCCACNAGCAGGTACCGTCGATACCGGGGTGTGTGC	2650847
Qy	105	SerArgCysGlnGlyLysHisArg-----	112

TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:

Pred. No.: 0.899 Length: 4411529
Score: 152.50 Matches: 79
Percent Similarity: 30.54% Conservative: 45
Best Local Similarity: 19.46% Mismatches: 100
Query Match: 10.64% Indels: 183
DB: 3 Gaps: 14

US-10-049-742-11 (1-269) x US-09-103-840A-1 (1-4411529)

```
QY 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
   :::::
Db 2653970 TAGCGCTCTCGCGGTGAGCAAGACGACGACGACATCAAAAGCGCTACCGC 2653911
   :::::

QY 32 GlnLeuAlaValMetValHisProAspLysAsnHisProArgAlaGluGluAlaPhe 51
   :::::
Db 2653910 AAGCTGCGCGGAGCTGATCCGACGTCAACCCGAGCAGGCTCGCAGGGGCAAAATTC 2653851
   :::::

QY 52 LysValLeuArgAlaAlaThrAspIleValSerAsnAlaGluLysArgLysGluTyrGlu 71
   :::::
Db 2653850 AAAGAAATCAGCGTCGCTACGAGGTGCTCAGTGACCCGACCAACGTCGATCGTCGAC 2653791
   :::::

QY 72 MetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPhe ----- 87
   :::::
Db 2653790 CTGGCGGGGATCCGCTGGAGAGCGCGCTCGCGGGCGCAATGGGTTCGGTTCGGC 2653731
   :::::

QY 87 ----- 87

Db 2653730 GGCCTCGGCGACGTGTTTCAGAGCGTTCCTTTGGCGGGGTTTCGGTGGGGCGCGCGCTCC 2653671

QY 88 ----- LeuSerLysLeuGlnAspAsp 94
   :::::
Db 2653670 CGAGGGCCGATCGGCGGGTCCGGCGGGTTCGGACTCGCTGCTAGAAATGGCGTGGAT 2653611
   :::::

QY 95 LeuLysGluAlaMet ----- AsnThrMetMetCys 104
   :::::
Db 2653610 CTCGAAGAGTGGCGCAACAGGTGTCAACAGAGGTACCGTTCGATACCCGCGGTGTGTGC 2653551
   :::::

QY 105 SerArgCysGlnGlyLysHisArg ----- 112
   :::::
Db 2653550 GACCGGTGCCAGGCAAGGGCCACCAACGGCGATTTCGGTTCGATACCTCGCACACCTGC 2653491
   :::::

QY 113 ----- ArgPheGluMetAspArgGluProLysSer ----- Ala 123
   :::::
Db 2653490 GGTGCCCGCGGGAGGTGCAGACCCGTGCAGCGATCGCTGTTGGTTCAGATGTTGACGTGC 2653431
   :::::

QY 124 ArgTyrCysAlaGluCys ----- 129
   :::::
Db 2653430 CGGCGGTGTCCACCTGCGCGGGTTCGGGTGTTATCCCGACCCGTGCCAGCAATGC 2653371
   :::::

QY 130 ----- AsnArgLeuHisProAlaGluGlu ----- 137
   :::::
Db 2653370 ATGGCGGATGGCGGATCCGGGCCCGTCCGGAGATCAGCGTCAGATCCGCCCGGTGTC 2653311
   :::::

QY 137 ----- 137

Db 2653310 GCGCAGGGATCGAGTTCGGTTCGCGCTCAGGGCAGGTCCGGCCCGGGGGAGGCGCG 2653251

QY 138 --- GlyAspPheTrpAlaGlu ----- 143
   :::::
Db 2653250 CGGGGTGACCTCTACGTGAGGTCCATGAGCAGGCGCACAGCTCTTTGTCGCGAAAGGT 2653191
   :::::

QY 144 ----- SerSerMetLeuGlyLeuLysIle 151
   :::::
Db 2653190 GACCATCTGCATTGCAGGTTTCGGTGGCCATGTTGACGCGCGCTGGGTGTACGGTC 2653131
   :::::

QY 152 ThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAlaGlyCysGln 171
   :::::
```

```
Db 2653130 ACGGTGGACGCCATCCTGGAC ----- 2653110

QY 172 ArgValGlyIleSerProAspThrHisArgValProTyrHisIleSerPheGlySerArg 191
   :::::
Db 2653109 -----GGCTTGAGCGA-GATCACCATTCCACCCGCGACGACGCCAGGTTTCGGTGTATC 2653057
   :::::

QY 192 IleProGlyThrArgGlyArgGlnAlaThrProAspAlaProProAlaAspLeuGln 211
   :::::
Db 2653056 GCTGCGCGGTTCGAGG -----AATCCCGCACCTGCG ----- 2653027
   :::::

QY 212 AspPheLeuSerArgIlePheGlnValPro -----ProGlyGlnMetProAsnGlyAsn 229
   :::::
Db 2653026 -----TTCCAAACACGCGTGGCGGACCTTCACGTTTCACCTGGAGGT --- 2652988
   :::::

QY 230 PhePheAlaAlaProGlnProAla -----ProGly -----AlaAlaAlaAla 243
   :::::
Db 2652987 -----GGTGGTCCCGACCGGCTGGATCACCAGGACATCGAACTGTCTGCGCGAGCTGAA 2652934
   :::::

QY 244 SerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArgLysVal 263
   :::::
Db 2652933 GGTCTCCCGCGACCGCGAGGTGGCGGAGTCCGTTTCGACCCACGCGCGCGCGGACT 2652874
   :::::

QY 264 ArgArgProPheGlnArg 269
   :::::
Db 2652873 GTTCAGCCCGGTTGCGCGA 2652856
   :::::
```

Search completed: December 1, 2003, 12:18:31
Job time : 3389 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 1, 2003, 10:10:04 ; Search time 362 Seconds
(without alignments)
2448.264 Million cell updates/sec

Title: US-10-049-742-11
Perfect score: 1433
Sequence: 1 MAGVPEDELNPFHVLGVZAT.....VPKGEAKPKRRKKVRRPQR 269

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delcp 6.0 , Delcxt 7.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+pn.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10049742/runat_01122003_073138_18649/app_query.fasta_1.455
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MOD=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10049742 -GCN 1.1_350 @runat_01122003_073138_18649
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygaext=0.5 -DELOP=6 -DELEXT=7

Database :		Published Applications NA:*	
		1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
		2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
		3:	/cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
		4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
		5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
		6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
		7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
		8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
		9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
		10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
		11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
		12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
		13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
		14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
		15:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
		16:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
		17:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
SUMMARIES			
Result	Query	Description	
No.	Score	Match Length	DB ID

c 1	1433	100.0	2945 14 US-10-198-846-12415 Sequence 12415, A

ALIGNMENTS	
RESULT 1	
US-10-198-846-12415/c	
; Sequence 12415, Application US/10198846	
; Publication No. US2003009974A1	
; GENERAL INFORMATION:	
; APPLICANT: Lillie, James	
; APPLICANT: Xu, Yongyao	
; APPLICANT: Wang, Youzhen	
; APPLICANT: Steinmann, Kathleen	
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS	
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND	
; TITLE OF INVENTION: THERAPY OF BREAST CANCER	
; FILE REFERENCE: MRI-049	
; CURRENT APPLICATION NUMBER: US/10198, 846	
; CURRENT FILING DATE: 2002-07-18	
; PRIOR APPLICATION NUMBER: 60/306,220	
; PRIOR FILING DATE: 2001-07-18	
; NUMBER OF SEQ ID NOS: 14084	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 12415	
; LENGTH: 2945	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-10-198-846-12415	

Alignment Scores:

Pred. No.: 2,71e-147 Length: 2945
Score: 1433.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-049-742-11 (1-269) x US-10-198-846-12415 (1-2945)

QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
Db 1576 ATGGCTGGGGTTCCTGAGGATGAGCTAAACCCCTTCCATGATCTGGGGTGTAGGCCACA 1517

QY 21 AlaSerAspValGluLeuLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
Db 1516 GCATCAGATGTTGAACGAAGAGGCTATAGACAGCTGCAGTGTGTTTCATCTGAC 1457

QY 41 LysAsnHisHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
Db 1456 AAAATCATCATCCCCGGGCTGAGGAGGCTTCAAGGTTTTCGAGCAGCTTGGGACATT 1397

QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
Db 1396 GTACGCAATGCTGAAAAGCGAAAGAGTATGAGATGAACGATGCGCAGAGATGAGCTG 1337

QY 81 SerAsgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
Db 1336 AGCCGGTCAGTAATGAGTTCTGTCGAAGCTGCAAGATGACCTCAAGGAGCAATGAAT 1277

QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluPro 120
Db 1276 ACTATGATGTGAGCGGATGCCAAGGAAAGCATGAGGAGTTTGAATGACCGGAACT 1217

QY 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPhe 140
Db 1216 AAGATGCCAGATACGTGCTGAGTGAATAGCTGCATCTCTGCTGAGGAAGAGACTTT 1157

QY 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 1156 TGGGCAGAGTCAGCATGTTGGCCCTCAAGATCACCTACTTTGCATGATGATGGAAG 1097

QY 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
Db 1096 GTGTATGACATCACAGAGTGGGCTGGATGCCAGCGTGTAGGTATCTCCCCAGATACCCAC 1037

QY 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
Db 1036 AGAGTCCCTCATCATCTCATTTGTTCTCGGATTCAGGCCAGACCGGCGGAGAGA 977

QY 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
Db 976 GCCACCCAGATGCCCTCTCTGCTGATCTTCAGGATTTCTGAGTCGGATCTTTCAAGTA 917

QY 221 ProProGlyGlnMetProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAla 240
Db 916 CCCCCAGGCGAGATGCCAATGGAACTTCTTTGAGCTCTCTCAGCTCCCTCAGCCTGCCCTCGAGCC 857

QY 241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
Db 856 GCTGAGGCTCTTAAAGCCCAACAGCAGTACCCAAAGGGAGAGCCAAACCTTAAGCGGCGG 797

QY 261 LysLysValArgArgProPheGlnArg 269
Db 796 AAGAAAGTGAGGAGGCCCTTCCAAAGT 770

RESULT 2

US-09-814-353-19835/c
; Sequence 19835, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela

APPLICANT: Lillie, James
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF OVARIAN CANCER
FILE REFERENCE: MRI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/191,031
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19835
LENGTH: 3090
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 488..525
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19835

Alignment Scores:
Pred. No.: 2.88e-147 Length: 3090
Score: 1433.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-049-742-11 (1-269) x US-09-814-353-19835 (1-3090)

QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
Db 1570 ATGGCTGGGGTTCCTGAGGATGAGCTAAACCCCTTCCATGATCTGGGGTGTAGGCCACA 1511

QY 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
Db 1510 GCATCAGATGTTGAACGAAGGCTATAGACAGCTGCAGTGTGTTTCATCTGAC 1451

QY 41 LysAsnHisHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
Db 1450 AAAATCATCATCCCCGGGCTGAGGAGGCTTCAAGGTTTTCGAGCAGCTTGGGACATT 1391

QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
Db 1390 GTACGCAATGCTGAAAAGCGAAAGGATGATGAGATGAACGATGCGCAGAGATGAGCTG 1331

QY 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
Db 1330 AGCCGGTCAGTAATGAGTTCTGTCGAAGCTGCAAGATGACCTCAAGGAGCAATGAAT 1271

QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluPro 120
Db 1270 ACTATGATGTGAGCCGATGCCAAGGAAAGCATAGGAGGTTTGAATGACCGGAACT 1211

QY 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPhe 140
Db 1210 AAGATGCCAGATACGTGCTGAGTGTATAGCTGTCATCTCTGCTGAGGAAGAGACTTT 1151

QY 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 1150 TGGGCAGAGTCAGCATGTTGGGCTTCAAGATCACCTACTTTCACCTGATGATGGAAG 1091


```
; ORGANISM: Homo sapiens
US-10-097-340-182

Alignment Scores:
Pred. No.: 4.38e-106 Length: 754
Score: 1051.00 Matches: 200
Percent Similarity: 99.50% Conservative: 0
Best Local Similarity: 99.50% Mismatches: 1
Query Match: 73.34% Indels: 1
DB: 14 Gaps: 0

US-10-049-742-11 (1-269) x US-10-097-340-182 (1-754)
QY 69 GluTyrGluMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeu 88
DB 2 GAGTATGAGTAAACAAATGGCAGAGAAATGAGCTGAGCCGGTCAGTAAATGAGTTCTG 61
QY 89 SerLysLeuGlnAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGln 108
DB 62 TCCAAGCTGCAAGATGACCTCAAGAGGCAATGAATATGATGTTAGCCGATGCCAA 121
QY 109 GlyLysHisArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGlu 128
DB 122 GGAAGCATAGGAGTTTGAATGGACCCGGGAACCTAAGAGTGCAGATATCTGTCTGAG 181
QY 129 CysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGly 148
DB 182 TGAAATAGGCTGCATCTGCTGAGGAAGGAGACTTTTGGCAGAGTCAAGCATGTTGGGC 241
QY 149 LeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAla 168
DB 242 CTCAGATACCTACTTTGACATGATGGATGGAAAGGTGTATGACATCACAGAGTGGCT 301
QY 169 GlyCysGlnArgValGlyIleSerProAspThrHisArgValProTyrHisIleSerPhe 188
DB 302 GGATGCCAGCGTGTAGTATCTCCACAGATACCCACAGAGTCCCTATCATCTCATTT 361
QY 189 GlySerArgIleProGlyThrArgGlyArgGlnArgAlaThrProAspAlaProProAla 208
DB 362 GGTCTCGGATTCAGGACACAGAGGGCGGACAGAGACCCACAGATGCCCTCTCTGCT 421
QY 209 AspLeuGlnAspPheLeuSerArgIlePheGlnValProGlyGlnMetProAsnGly 228
DB 422 GATCTTCAGGATTTCTTGTAGTCGATCTTCAAGTACCCCGGCGAGATGCCAAT-GGG 480
QY 229 AsnPhePheAlaProGlnProAlaProGlyAlaAlaAlaSerLysProAsnSer 248
DB 481 AACTTCTTTCAGCTCTCTAGCTGCCCCCTGGAGCCGCTGCAGCCTCTAAGCCCAACAGC 540
QY 249 ThrValProLysGlyGluAlaLysProLysArgArgLysLysValArgArgProPheGln 268
DB 541 ACAGTACCCCAAGGAGAGCAAAACCTTAAGCGGGAAGAAAGTGAGGAGGCCCTTCCAA 600
QY 269 Arg 269
DB 601 CGT 603

RESULT 5
US-09-764-891-8949
; Sequence 8949, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR APPLICATION DATA: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8949
; LENGTH: 7453
; TYPE: DNA
```


Db 4772 CTTGTCACATCTCATTTTGGTTCACGGATGCCT 4804

```

RESULT 9
US-10-228-406A-9
; Sequence 9, Application US/10228406A
; Publication NO. US20030104612A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; APPLICANT: Zybarrh, Gabriele
; TITLE OF INVENTION: GENERATION OF TYPE I/TYPE II HYBRID FORM OF BOVINE
; TITLE OF INVENTION: VIRAL DIARRHEA VIRUS FOR USE AS VACCINE
; FILE REFERENCE: PC11051A
; CURRENT APPLICATION NUMBER: US/10/228,406A
; CURRENT FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 9
; LENGTH: 16713
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pNADL890 vector
US-10-228-406A-9

```

100-10-328 4063 0

```

US-10-228-406A-9
Alignment Scores:
Pred. No.:      1.84e-43      16713
Score:          497.00      Matches:      89
Percent Similarity: 98.90%      Conservative: 1
Best Local Similarity: 97.80%      Mismatches: 1
Query Match:     34.68%      Indels:      0
DB:              14          Gaps:        0
US-10-049-742-11 (1-269) x US-10-228-406A-9 (1-16713)

```

Db	5108	GAGTCGAGCATGTTGGCGCCTCAAAATCACCTACTTTGCGCTGATGTGATGAAAGGTGAT	516
Qy	163	AspIleThrClnrTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgVal	182
Db	5168	GATATCACAGATGGCGTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCACAGATC	5227
Qy	183	ProTyrHisIleSerPheGlySerArgIlePro	193
Db	5228	CCTGTGCATCATCTCATTTGGTTCCACGGATGCCT	5260

D**b** 5228 CCTGTCACATCTCATTTGGTTCACGGATGCCT 5260

RESULT 10
US-09-814-353-1424/C
; Sequence 1424, Application US/09814353
; Publication No. US20030165931A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940

```

; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1424
; TYPE: DNA
; LENGTH: 259
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2_
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-1424

Alignment Scores:
Pred. No.: 2,58e-40 Length: 259
Score: 447.00 Matches: 85
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.19% Indels: 0
DB: 12 Gaps: 0

US-10-049-742-11 (1-269) x US-09-814-353-1424 (1-259)
QY 30 TyrArgGlnLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGlu 49
DB 258 TATAGACAGCTGGCAGTGATGTTTCATCTCTGACAAAATCATCATCCCCGGCTGAGGAG 199
QY 50 AlaPheLysValLeuArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGlu 69
DB 198 GCCTTCAAGTTTTGCGAGCACCTTGGGACATTGTGACGAATGCTGAAAAGCGAAGGAG 139
QY 70 TyrGluMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSer 89
DB 138 TATGAGATGAACAGCAATGCGACAGATGAGCTGAGCGGTCAGTAAATGAGTTTCTGTC 79
QY 90 LysLeuGlnAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGlnGly 109
DB 78 AAGCTGCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTGTAGCCGATGCCAAGGA 19

QY 110 LysHisArgArgPhe 114
DB 18 AAGCATAGAGGTTT 4

RESULT 11
US-09-814-353-7786/c
; Sequence 7786, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25

```

```

; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7786
; LENGTH: 259
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2_
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-7786

Alignment Scores:
Pred. No.: 2,58e-40 Length: 259
Score: 447.00 Matches: 85
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.19% Indels: 0
DB: 12 Gaps: 0

US-10-049-742-11 (1-269) x US-09-814-353-7786 (1-259)
QY 30 TyrArgGlnLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGlu 49
DB 258 TATAGACAGCTGGCAGTGATGTTTCATCTCTGACAAAATCATCATCCCCGGCTGAGGAG 199
QY 50 AlaPheLysValLeuArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGlu 69
DB 198 GCCTTCAAGTTTTGCGAGCACCTTGGGACATTGTGACGAATGCTGAAAAGCGAAGGAG 139
QY 70 TyrGluMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSer 89
DB 138 TATGAGATGAACAGCAATGCGACAGATGAGCTGAGCGGTCAGTAAATGAGTTTCTGTC 79
QY 90 LysLeuGlnAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGlnGly 109
DB 78 AAGCTGCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTGTAGCCGATGCCAAGGA 19

QY 110 LysHisArgArgPhe 114
DB 18 AAGCATAGAGGTTT 4

RESULT 12
US-09-764-868-175
; Sequence 175, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 1030
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (776)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (923)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1002)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-175

```

Alignment Scores: 6,69e-12 Length: 1030
Pred. No.: 74
Score: 195.50 Matches: 74
Percent Similarity: 39.10% Conservative: 30
Best Local Similarity: 27.82% Mismatches: 78
Query Match: 13.64% Indels: 84
DB: 10 Gaps: 11

US-10-049-742-11 (1-269) x US-09-764-868-175 (1-1030)

Qy 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
Db 381 TATGAGATCTCTGGGGTGGAGAGAGGGCTCGGATGAGGACCTGAAGAGGCTTACCGC 440
Qy 32 GinLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPhe 51
Db 441 AGACTGGCCCTCAAAATTCACCCAGACAGAACCCAGCACCCTGGTCCACTGAAGCCTTC 500
Qy 52 LysValLeuArgAlaAlaTrpAspLeuValSerAsnAlaGluLysArgLysGluTyrGlu 71
Db 501 AAAGCCATTGGCACAGCATATCGGTACTACGAAACCCGAGAGAGGAGGATGATGAC 560
Qy 72 MetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeu 91
Db 561 -----CAGTTC 566
Qy 92 GinAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGlnGlyLys 110
Db 567 GCGGATGACAGAGCCAGCGCCCGG-----CACGGCCATGGG 605
Qy 111 -----HisArgArgPheGluMetAspArgGluPro-LysSerAlaArgTyrCy 126
Db 606 CATGGGGATTTCCACCGTGGCTTTGAGGCGGACATCTCCCTCGAAGACCTCTTCAACATG 665
Qy 126 sAlaGluCysAsnArgLeuHisProAlaGluGlyAspPheTrpAlaGluSerSerMe 146
Db 666 TTCT-----TTGGCGCGGCTTCCCT 686
Qy 146 tLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGl 166
Db 687 TCTAGTAACGTCACGCTCT-----ArgValGlyIleSerPro----- 177
Qy 166 utrpAlaGlyCysGln-----CACGGCCATGGG 605
Db 711 AACGGCCGATCGGTATACCTACACAGCAAGAGGACCGGAGCAAAACCGGGTGATG 770
Qy 178 -----AspThrHisArgValPro---TyrHisIleSerPh 188
Db 771 GCGGNTAGGGGTGTTGTGCAGTGTATCGCTATCTCTGATTCCTGTTGTCAGCTT 830
Qy 188 eGlySerArgIleProGly-----ThrArgGlyArgGlnArgAlaThrProAspAl 205
Db 831 CTTACGCCAGCTTCATGCTCCAGTCCACCTACAGTCTGAGTCCAGACCCGCTCGTG 890
Qy 205 aProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnValProProGlyGlnMe 225
Db 891 GCCACATCCAGGCGAGTCAATGACCACTNGGTGTTCTGCTTATTATGTTGGGAG-- 948
Qy 225 tProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerLys 245
Db 949 -----GACATTTTCCAGGGTTACAAGGTTTCCA-----GCTTCAAA 989
Qy 245 sProAsnSerThrVal 250
Db 990 ACATTCGAGCGGNATT 1005

RESULT 13
US-09-764-868-548
; Sequence 548, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 548
LENGTH: 1030
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (776)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (923)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1002)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-548

Alignment Scores: 6,69e-12 Length: 1030
Pred. No.: 74
Score: 195.50 Matches: 74
Percent Similarity: 39.10% Conservative: 30
Best Local Similarity: 27.82% Mismatches: 78
Query Match: 13.64% Indels: 84
DB: 10 Gaps: 11

US-10-049-742-11 (1-269) x US-09-764-868-548 (1-1030)

Qy 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyrArg 31
Db 381 TATGAGATCTCTGGGGTGGAGAGAGGGCTCGGATGAGGACCTGAAGAGGCTTACCGC 440
Qy 32 GinLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPhe 51
Db 441 AGACTGGCCCTCAAAATTCACCCAGACAGAACCCAGCACCCTGGTCCACTGAAGCCTTC 500
Qy 52 LysValLeuArgAlaAlaTrpAspLeuValSerAsnAlaGluLysArgLysGluTyrGlu 71
Db 501 AAAGCCATTGGCACAGCATATCGGTACTACGAAACCCGAGAGAGGAGGATGATGAC 560
Qy 72 MetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeu 91
Db 561 -----CAGTTC 566
Qy 92 GinAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGlnGlyLys 110
Db 567 GCGGATGACAGAGCCAGCGCCCGG-----CACGGCCATGGG 605
Qy 111 -----HisArgArgPheGluMetAspArgGluPro-LysSerAlaArgTyrCy 126
Db 606 CATGGGGATTTCCACCGTGGCTTTGAGGCGGACATCTCCCTCGAAGACCTCTTCAACATG 665
Qy 126 sAlaGluCysAsnArgLeuHisProAlaGluGlyAspPheTrpAlaGluSerSerMe 146
Db 666 TTCT-----TTGGCGCGGCTTCCCT 686
Qy 146 tLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGl 166
Db 687 TCTAGTAACGTCACGCTCT-----ArgValGlyIleSerPro----- 177
Qy 166 utrpAlaGlyCysGln-----CACGGCCATGGG 605
Db 711 AACGGCCGATCGGTATACCTACACAGCAAGAGGACCGGAGCAAAACCGGGTGATG 770
Qy 178 -----AspThrHisArgValPro---TyrHisIleSerPh 188
Db 771 GCGGNTAGGGGTGTTGTGCAGTGTATCGCTATCTCTGATTCCTGTTGTCAGCTT 830
Qy 188 eGlySerArgIleProGly-----ThrArgGlyArgGlnArgAlaThrProAspAl 205
Db 831 CTTACGCCAGCTTCATGCTCCAGTCCACCTACAGTCTGAGTCCAGACCCGCTCGTG 890
Qy 205 aProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnValProProGlyGlnMe 225
Db 891 GCCACATCCAGGCGAGTCAATGACCACTNGGTGTTCTGCTTATTATGTTGGGAG-- 948
Qy 225 tProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerLys 245
Db 949 -----GACATTTTCCAGGGTTACAAGGTTTCCA-----GCTTCAAA 989
Qy 245 sProAsnSerThrVal 250
Db 990 ACATTCGAGCGGNATT 1005

Db 831 CTTACGACGAGTTCATGCTCCAGTCCACCTCAGTCTGAGTCCAGACCGTCCGTGG 890
Qy 205 aProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnValProGlyGlnMe 225
Db 891 GCCACATCCACGCGAGTCAATGACCACTGCTGCTTATTATGTTGGGAG-- 948
Qy 225 tProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerIly 245
Db 949 -----GACAATTTTCCAGGGTTCAAAAGGTTTCCA-----GCTTCAA 989
Qy 245 sProAsnSerThrVal 250
Db 990 ACATTCGAGCGGNATT 1005
RESULT 14
US-10-252-157-105
; Sequence 105, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 105
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 401434.10
US-10-252-157-105

Alignment Scores:
Pred. No.: 1.4e-10 Length: 3025
Score: 189.00 Matches: 44
Percent Similarity: 50.43% Conservative: 14
Best Local Similarity: 38.26% Mismatches: 27
Query Match: 13.19% Indels: 30
DB: 12 Gaps: 3
US-10-049-742-11 (1-269) x US-10-252-157-105 (1-3025)
Qy 12 PheHisValLeuGlyValGluAlaThrAlaSerAspValGluLeuLysLysAlaTyArg 31
Db 416 TATGAGATCCTGGGGGTGAGCAGAGGGGCTCGGATGAGGACCTGAAGAAGCCCTACCGC 475
Qy 32 GlnLeuAlaValMetValHisProAspLysAsnHisHisProArgAlaGluGluAlaPhe 51
Db 476 AGACTGGCCCTCAATTCACCCAGACAGAACACCGACCTGGTGCACCTGAAGCCCTTC 535
Qy 52 LysValLeuArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyArg 71
Db 536 AAAGCATTGGCACAGCATATCGGTACTACACCAACCCGAGAGAGGAGGAGTATGAC 595
Qy 72 MetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeu 91
Db 596 -----CAGTTC 601
Qy 92 GlnAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGlnLys--- 110
Db 602 GCGGATGACAAAGAGCGGCGGCCGCG-----CACGGCCATGGG 640
Qy 111 -----HisArgArgPheGluMetAspArgGluProLys 121
Db 641 CATGGGGATTCCACCGCTGGCTTTGAGGCCGACATCTCCCTGAA 685

RESULT 15
US-10-198-846-5395
; Sequence 5395, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5395
; LENGTH: 928
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 413, 430, 468, 500, 519, 521, 537, 546, 552, 554, 559, 571,
; LOCATION: 581, 586, 589, 592, 608, 615, 640, 644, 652, 653, 664, 676,
; LOCATION: 679, 685, 690, 695, 699, 706, 707, 721, 722, 725, 728, 730,
; LOCATION: 740, 742, 754, 761, 765, 769, 775, 777, 794, 807, 813
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 817, 819, 831, 833, 837, 840, 845, 851, 852, 858, 860, 861,
; LOCATION: 864, 865, 867, 868, 877, 878, 879, 880, 881, 884, 885, 897,
; LOCATION: 899, 900, 905, 912, 913, 918, 919, 920, 924, 927, 928
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5395

Alignment Scores:
Pred. No.: 2.95e-10 Length: 928
Score: 180.00 Matches: 38
Percent Similarity: 60.24% Conservative: 12
Best Local Similarity: 45.78% Mismatches: 33
Query Match: 12.56% Indels: 0
DB: 14 Gaps: 0
US-10-049-742-11 (1-269) x US-10-198-846-5395 (1-928)
Qy 3 GlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThrAlaSer 22
Db 6 GGAGTCCACCGCGGTGGCG 65
Qy 23 AspValGluLeuLysLysAlaTyArgGlnLeuAlaValMetValHisProAspLysAsn 42
Db 66 GATGAGATTGAAAAAAGCTTATAGAAAGCTTGCTTTGAAGTTTCATCCAGACAAAAAC 125
Qy 43 HisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIleValSer 62
Db 126 CATGCACCTGGAGCAACAGATGCTTTAAAAAGATTGGAATGCTTATGCTGTTTAAAGT 185
Qy 63 AsnAlaGluLysArgLysGluTyArgGluMetLysArgMetAlaGluAsnGluLeuSerArg 82
Db 186 AATCCAGAAAACGAAACAGTATGACCTCAGCGGCAATGAAGAACAGCATGTAACCCAC 245
Qy 83 SerValAsn 85
Db 246 CAAAAAAT 254

Search completed: December 1, 2003, 12:12:32
Job time : 596 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 1, 2003, 09:38:38 ; Search time 2530 Seconds

(without alignments)
2584.152 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433

Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPKGEAKPKRKVKRRPFQR 269

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10049742/runat_01122003_073136_18615/app.query.fasta_1.455
-DB=EST -QPMI=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10049742 @CGN 1 1 2810 @runat_01122003_073136_18615 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1361	95.0	2300	11	AK017830 Mus muscu
2	1361	95.0	2639	11	AK032681 Mus muscu
3	1361	95.0	2712	11	AK045445 Mus muscu
4	1354	94.5	4231	11	AK082579 Mus muscu
5	1348	94.1	3022	11	AK045345 Mus muscu
6	1337	93.3	4053	11	AK083878 Mus muscu
7	1280	89.3	942	13	BU538214 AGENCOURT
8	1240	86.5	955	13	BU538351 AGENCOURT
9	1224	85.4	957	12	BM802135 AGENCOURT
10	1189	83.0	918	13	BU956189 AGENCOURT
11	1138	79.4	878	12	BU763500 602735591
12	1138	79.4	904	13	BQ691177 AGENCOURT
13	1076.5	75.1	723	10	BE779250 601464514
14	1076	75.1	927	12	BM424147 AGENCOURT
15	1034	72.2	874	10	BF679562 602153594
16	1025	71.5	737	14	CB324690 UI-R-D20-
17	1020	71.2	756	12	BG868350 602787804
18	1018	71.0	871	10	BF139916 601789432
19	1013	70.7	900	10	BG298053 602396271
20	1005.5	70.2	787	13	BU288529 603608051
21	989	69.0	679	9	AW146163 um38b12.x
22	973	67.9	785	12	BI333110 602996387
23	969	67.6	670	13	BU708368 UI-M-FCO-
24	964	67.3	667	14	BY760298 BY760298
25	946.5	66.1	675	13	BU407874 604139594
26	940	65.6	723	10	BE619407 601473229
27	930	64.9	560	9	AV602887 AV602887
28	927	64.7	715	13	BU323784 603493402
29	919	64.1	676	13	BU229462 603400604
30	887	61.9	651	10	BB465186 BB465186
31	864	60.3	645	13	BU371701 603811742
32	854	59.6	781	14	CB520550 UI-M-G10-
33	837	58.4	912	10	BG751088 602729879
34	835	58.3	810	13	BU247142 603551604
35	821	57.3	598	12	BM208252 C0624E11-
36	812	56.7	810	13	BU219192 603108510
37	809.5	56.5	527	12	BI002288 PM1-HN007
38	796	55.5	618	12	BM011379 603635574
39	796	55.5	857	13	BU900250 AGENCOURT
40	794	55.4	481	10	BF722658 mab34b06
41	794	55.4	605	13	EX090711 EX090711
42	793	55.3	467	10	BE864884 UI-M-BH1-
43	786	54.8	448	12	BI002295 PM1-HN007
44	785	54.8	683	13	BU287112 604166895
45	783	54.6	453	9	AI504202 vk90d12.x

ALIGNMENTS

RESULT 1	AK017830	2300 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS	AK017830	Mus musculus 8 days embryo whole body cDNA, RIKEN full-length			
DEFINITION		enriched library, clone:570551F12 product:DOPAMINE RECEPTOR			
		INTERACTING PROTEIN homolog [Rattus norvegicus], full insert			
		sequence.			
ACCESSION	AK017830				
VERSION	AK017830.1	GI:12857283			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

Qy	21	AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProasp	40	REFERENCE
Db	1459	GCATCCGACACTGAACCTAAAGAAGCCCTATAGCGAGCTAGCAGTAAATGCTCCATCTGTAT	1518	AUTHORS
Qy	41	LysAsnHisHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTrpaspIle	60	TITLE
Db	1519	AAAAATCAACCTCCCGCGCTGAGGAGGCTTTCAAAATTTTTCGGGCGAGCTTGGGACATT	1578	JOURNAL MEDLINE PUBMED
Qy	61	ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu	80	REFERENCE
Db	1579	GTCAGCAACCCAGAGAGCGGGAAGAAATATGATGAACCGATGGCAGAGAAATGAGTCT	1638	AUTHORS
Qy	81	SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn	100	TITLE
Db	1639	AGCGGTGAGTGAATGATGTTCTGTGAAACTACAGATGACCTCAGGAGGCAATGAC	1698	JOURNAL MEDLINE PUBMED
Qy	101	ThrMetMetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluPro	120	REFERENCE
Db	1699	ACGATGATGTGCAGCAGATGCCAAGAAAGCATAGGAGGTTTGAATGGACCGGAACCC	1758	AUTHORS
Qy	121	LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGlyAspPhe	140	TITLE
Db	1759	AAGAGTGCAGATAGTGTGCTGAGTGTAAACAGGCTGCATCTCTGCTCAGGAAGGACATTT	1818	JOURNAL MEDLINE PUBMED
Qy	141	TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys	160	REFERENCE
Db	1819	TGGGAGAGTGCAGATGCTGGGCTCAAGATCACTTACTTTGCCGTGATGATGGAAG	1878	AUTHORS
Qy	161	ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis	180	TITLE
Db	1879	GTGATGACATCACAGATGGGCTGATGCCAGCGTGTGGTATCTCCCGATCTACTAC	1938	JOURNAL MEDLINE PUBMED
Qy	181	ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg	200	REFERENCE
Db	1939	AGAGTTCTTACCACATCTCATTTGTTCTCGGTACCGGCACCGAGTGGCCGCGAGAGG	1998	AUTHORS
Qy	201	AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal	220	TITLE
Db	1999	CCCACTCCAGATGCCCTCTGCTGACCTGCGAGGATTTCTTGAGCGGATCTTTCAAGTA	2058	JOURNAL MEDLINE PUBMED
Qy	221	ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla	240	REFERENCE
Db	2059	CCTCGGGGCGGATGCAATGGGAACCTCTTTGCGGCACCTCACCTGGCCCTGGGACC	2118	AUTHORS
Qy	241	AlaAlaLaserLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg	260	TITLE
Db	2119	ACTTCGACCTTAGGCCCAACAGTTTCAGTACCAGGCGGAGGCAACCTTAACGGCGG	2178	JOURNAL MEDLINE PUBMED
Qy	261	LysLysValArgArgProPheGlnArg	269	REFERENCE
Db	2179	AAGAAAGTGAGGCGGCCCTTTCAACGA	2205	AUTHORS
RESULT 2				
AK032681				
LOCUS				
DEFINITION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				

le Dec 2 08:21:31 2003

Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

NURS

source
1. .2639
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:6530416L18"
/db_xref="taxon:10090"
/clone="6530416L18"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
64. -2176
/note="DOPAMINE RECEPTOR INTERACTING PROTEIN homolog
[Rattus norvegicus] (SPTR|Q925G7, evidence: FASTV,
92.7%ID, 100%length, match=2109)
putative"

misc_feature

666 a 634 c 738 g 601 t
putative

SE COUNT
IGIN

ignment Scores: Length: 2639
ed. No.: 1.47e-107
ore: 1361.00
Matches: 251
Conservative: 8
Mismatch: 10
Indels: 0
Gaps: 0
Percent Similarity: 93.31%
Local Similarity: 94.98%
Very Match: 11

B:
S-10-049-742-11 (1-269) x AK032681 (1-2639)

1 MetAlaGlyValProGluAspGluLeuAsnPropHeHisValLeuGlyValGluAlaThr 20
1367 ATGGCTGGGGTTCCTGAAGATGAATTAACCTTTTCATGTGCTGGGGTTGAAGCTACA 1426
21 AlaSerAspValGluLeuLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
1427 GCATCCGACACTGAATTAAGAGGCTATAGCAGCTAGCATGATGATGCTCATCTGAT 1486
41 LysAsnHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60
1487 AAAATCACCATCCCGGGCTGAGGAGGCTTCAAAATTTTGGGCGAGCTTGGGACATT 1546
61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
1547 GTCAGCAACCCAGAGAGCGGGAAGGAATATGAGATGAACGATGGCAGAGATGAGCTC 1606
81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsn 100
1607 AGCCGGTCAGTGAATGAGTTTCTGCGAACTACAAAGATGACCTCAAGAGGCGATGAC 1666
101 ThrMetMetCysSerArgCysGlnGlyLysHisArgAtgPheGluMetAspArgGluPro 120
1667 ACGATGATGTCAGCAGATGTCGAAGGAAAGCATGAGGATTTGAATGACCGGGAACCC 1726
121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluArgPhe 140
1727 AAGAGTCCAGATGATGCTGTGATGATGAACGCTGCATCTCTGAGGAAGGAGACTTT 1786
141 TrpAlaGluSerSerMetLeuGlyLysLysIleThrTyrPheAlaLeuMetAspGlyLys 160
1787 TGGCAGATGCGAGATGCTGGGCTCAAGATCACTTACCTTGGCTGCTGATGATGGAAG 1846
161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
1847 GTGTATGACATCAGAGATGGGCTGGATGCCAGCGTGTGGGTATCTCCCGAGATCTCAC 1906
200

1907 AGATTCTTACCACATCTCATTTGGTTCTCGGTACCCGCGCACCATGTCGCGGAGG 1966
201 AlathrProAspAlaProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
1967 GCCACTCCAGAGTCCCTCTCTGCTGACCTGCGAGGATTTCTTGACCGGATCTTTCAAGTA 2026
221 ProProGlyGlnMetProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAla 240
2027 CTTCCGGGGCGGATGTCCATGGAACCTTTTGGCGACCTCACCTGGCCCTGGGACC 2086
241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
2087 ACTTCGACCTTAGCGCCACAGTTCAGTACCCCAAGGAGGAGGAGCAACCTTAAAGGCGG 2146
261 LysLysValArgArgProPheGlnArg 269
2147 AAGAAAGTGAGCGCGCTTTCAACGA 2173

RESULT 3

AK045445
LOCUS
DEFINITION

Mus musculus adult male corpora quadrigemina cDNA, RIKEN
full-length enriched library, clone:B230120106 product:DOPAMINE
RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus], full
insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUB ID

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED

Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Meshima, J., Mazzei, J., Momberte, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyoh-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
MEDLINE
11217851

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)
6 (bases 1 to 2712)

Adachi, J., Alzawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers
1. 2712
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B230120106"
/db_xref="taxon:10090"
/clone="B230120106"
/sex="male"
/tissue_type="corpora quadrigemina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
177. 2288
/note="unnamed protein product: DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus] (SPR|Q925G7, evidence: FASTY, 92.%ID, 100%length, match=2109)
putative"
/codon_start=1
/protein_id="BAC123172.1"
/db_xref="GI:26337373"
/translation="MAQKHGERRUCGAHRSQGTSLTSGSVDPFELISFGLRDSAE
TAPNRTCKSHGSKYQTPNPAPWSPDPSHPGPPGPGPPGPGGYPDSSETGSESGV
DOELSRNETGVQEDSGSFLSIIPACNCQSPGVPGETYSEGGSSSLCHCTSP
ALGDELEEDYDDPELPKPSDRSVSSGKKPLSRKROKREFLIKEDYDSGERPKA
PGRRLAKRSQTDKRGILNGVZELQCLGQAGFWLIELLVINGEVETCGHLIYA
CRKLKGLDLDVFWVWVGWARRLGGWARMFQLSQSFVGVGLLIRLVVVGAFLLI

CDs
177. 2288
/note="unnamed protein product: DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus] (SPR|Q925G7, evidence: FASTY, 92.%ID, 100%length, match=2109)
putative"
/codon_start=1
/protein_id="BAC123172.1"
/db_xref="GI:26337373"
/translation="MAQKHGERRUCGAHRSQGTSLTSGSVDPFELISFGLRDSAE
TAPNRTCKSHGSKYQTPNPAPWSPDPSHPGPPGPGPPGPGGYPDSSETGSESGV
DOELSRNETGVQEDSGSFLSIIPACNCQSPGVPGETYSEGGSSSLCHCTSP
ALGDELEEDYDDPELPKPSDRSVSSGKKPLSRKROKREFLIKEDYDSGERPKA
PGRRLAKRSQTDKRGILNGVZELQCLGQAGFWLIELLVINGEVETCGHLIYA
CRKLKGLDLDVFWVWVGWARRLGGWARMFQLSQSFVGVGLLIRLVVVGAFLLI

ALALFLGCLQGLGMRFSVGLGNLGRWDRKTAMFLSMGLSPALHHLCLTLTKDSRPWOOLV
RLIQNGWELPWKQRTKQGNAPVSGRYCQPEBEVTRLLTMAGVPEDELNPFHLVG
VEATASDTLKKAYHQLAVLMDPKNNHPRBEAFKILRAADWDIVNPKRKEYEMKR
NAENLSRSVNEFLSKLDQDLKEAMNTWCSRQKRRFEMDRPKSARYCAECNRL
HPAEGSDFWAESMLGKITITYFALMDGKVIDITWAGCORVGTSPOTHRVYHISGGS
RVPVSGRQRATPESPPADLQDLSRIFQVPPGPMNSGNFVFAAPHEPGPGTTSTRPNS
SVFPGKAPKRRKKYRRFPQR"

BASE COUNT 664 a 663 c 778 g 607 t
ORIGIN

Alignment Scores:
Pred. No.: 151e-107 Length: 2712
Score: 1361.00 Matches: 251
Percent Similarity: 96.28% Conservative: 8
Best Local Similarity: 93.31% Mismatches: 10
Query Match: 94.98% Indels: 0
DB: 11 Gaps: 0

US-10-049-742-11 (1-269) x AK045445 (1-2712)

QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20
DB 1479 ATGGCTGGGGTCTCTGAAGATGAACCTAAACCTTTTCATGTGCTGGGGTTGAGCTACA 1538
QY 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40
DB 1539 GCATCCGACACTGAACATAAGAGGCTATAGGCAGCTAGCAGTAATGGTCCATCTCGAT 1598
QY 41 LysAsnHisHisProArgAlaGluAlaPheLysValLeuArgAlaAlaTyrAspIle 60
DB 1599 AAAAAATCAACATCCCGGGCTGAGGAGGCTTCAAAATTTTGGCGGAGCTTGGGACATT 1658
QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80
DB 1659 GTCAGCAACCCAGAGAGCGGAGGAAGTAATAGATGAATAAAGCGGATGCGAGAGATGAGCTC 1718
QY 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsn 100
DB 1719 ACCGGTCTAGTGAATGAGTTTCTGTCGAACTACAAAGATGACCTCAAGAGGCAATGAAC 1778
QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120
DB 1779 ACGATGATGTGAGCAGATGCCAAGAAAGCATAGGAGGTTTGAATGGACCGGAAACCC 1838
QY 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
DB 1839 AAGAGTGCCAGATACTGTCTGAGTGAACAGGCTCATCTCTCTGAGGAAGGAGACTTT 1898
QY 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
DB 1899 TGGGAGAGTGTGAGCATGCTGGGCTTCAAGATCACTTCTTTCGCTGATGGATGGAAG 1958
QY 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180
DB 1959 GTGTATGATCATCAGAGATGGGCTGATGCCAGCGTGTGGTATCTCTCCAGATCTAC 2018
QY 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
DB 2019 AGAGTCTCTTACCACATCTCAATTTGGTCTCGGATCCCGGACCCAGTGGCCGCGAGG 2078
QY 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
DB 2079 GCCACTCCAGATGCCCTCTGCTGACCTGCGAGGATTTCTTGGCGGATCTTTTCAAGTA 2138
QY 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240
DB 2139 CTTCGGGGCCCATGTCTCAATGGAACTTCTTTCGCGCACCTCACCTGGCCCTGGGACC 2198
QY 241 AlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
DB 2199 ACTTGACCTCTAGGCCCCCAACAGTTCAGTACCAAGGAGAGAACCACTTAAACGCGG 2258
QY 261 LysLysValArgArgProPheGlnArg 269

US-10-049-742-11 (1-269) x AK082579 (1-4231)		REFERENCE
QY	1 MetAlaGlyValProGluAspLeuLeuAanProPheHisValLeuGlyValGluAlaThr 20	AUTHORS TITLE JOURNAL MEDLINE PUBMED
Db	1386 ATGGCTGGGGTTCTCGAAGATGAATAAACCCCTTTTCATGTGCTGGGGTTGAAGCTACA 1445	REFERENCE
QY	21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40	AUTHORS
Db	1446 GCATCCGACCTGAACCTAAAGAGGCTATAGCGAGCTAGCAGTAAATGTGCTCATCTTGAT 1505	TITLE JOURNAL MEDLINE PUBMED
QY	41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIle 60	AUTHORS
Db	1506 AAAAATCACCATCCCGGGCTCAGGAGGCGCTTCAAAATTTTCGGGCGACCTGGGACATT 1565	TITLE JOURNAL MEDLINE PUBMED
QY	61 ValSerAsnAlaGluLysLysLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80	AUTHORS
Db	1566 GTCACCAACCCAGAGAGCGGGAAGAAATATGATGAAACCGATGCGAGAGATGAGCTC 1625	TITLE JOURNAL MEDLINE PUBMED
QY	81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsn 100	AUTHORS
Db	1626 AGCCGGTCAGTGAATGAGTTTCTGCGAACTACAGATGACCTCAAGGAGGCAATGAAC 1685	TITLE JOURNAL MEDLINE PUBMED
QY	101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120	AUTHORS
Db	1686 ACGATGATGTGCAGCAGATGCCAAGAAAGCATAGGAGGTTTGAATGGACCGGAACCC 1745	TITLE JOURNAL MEDLINE PUBMED
QY	121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140	AUTHORS
Db	1746 AAGAGTGGCGGATAGTGTCTGAGTGAACAGGCTGCATCTGCTGAGGAGGAGACATT 1805	TITLE JOURNAL MEDLINE PUBMED
QY	141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160	AUTHORS
Db	1806 TGGGCGAGTGCAGATCTGGCGCTCAAGATCACTTACTTTGCGCTGATGATGGAAG 1865	TITLE JOURNAL MEDLINE PUBMED
QY	161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180	AUTHORS
Db	1866 GTGTATGATACATCAGAGTGGCTGGATGCCAGCGTGGGTATCTCCCGATACTCAC 1925	TITLE JOURNAL MEDLINE PUBMED
QY	181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200	AUTHORS
Db	1926 AGAGTTCTTACCACATCTCATTTGGTCTCGGTACCCGACCCGAGTGGCCGGGAGAGG 1985	TITLE JOURNAL MEDLINE PUBMED
QY	201 AlaThrProAspAlaProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220	AUTHORS
Db	1986 GCCACTCCAGATGCCCTCTCTGCTCACCTGCAGGATTTCTTTCGAGCGGATCTTTCAGTA 2045	TITLE JOURNAL MEDLINE PUBMED
QY	221 ProProGlyGlnMetProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAla 240	AUTHORS
Db	2046 CCTCCGGGCGGATGTCCAATGGGAATCTTTTGGCGCACCTCACCTGGCCCTGGGACC 2105	TITLE JOURNAL MEDLINE PUBMED
QY	241 AlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260	AUTHORS
Db	2106 ACTTCGACCTCTAGGCCCAACAGTTTACGTACCAAGGAGAGCCAAACCTAAACGGCGG 2165	TITLE JOURNAL MEDLINE PUBMED
QY	261 LysLysValArgArgProPheGlnArg 269	AUTHORS
Db	2166 AAGAAAGTGAGCGGCCCTTTCAACGA 2192	TITLE JOURNAL MEDLINE PUBMED
RESULT 5	AK045345	
LOCUS	AK045345	
DEFINITION	Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230104C23 product:DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus], full insert sequence.	
ACCESSION	AK045345	
VERSION	AK045345.1 GI:26337270	
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1	Carninci, P. and Hayashizaki, Y.	
	High-efficiency full-length cDNA cloning	
	Metb. Enzymol. 303, 19-44 (1999)	
	99279253	
2	Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
	Genome Res. 10 (10), 1617-1630 (2000)	
	20499374	
3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, K., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hazada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
	Genome Res. 10 (11), 1757-1771 (2000)	
	20530913	
4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, K., Saito, T., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Resole, G., Tomita, M., Quackenbush, J., Schriml, L. M., Stauber, F., Suzuki, R., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyok, O., K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.	
	Functional annotation of a full-length mouse cDNA collection	
	Nature 409 (6821), 685-690 (2001)	
	21095660	
5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	
	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
	Nature 420, 563-573 (2002)	
	6 (bases 1 to 3022)	
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Kato, H., Kawai, J., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tsgami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	
	Direct Submission	
	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome	

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

Location/Qualifiers
1..3022
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B230104C23"
/db_xref="taxon:10090"
/clone="B230104C23"
/sex="male"
/tissue_type="corpora quadrigemina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
140..2176
/note="unnamed protein product; DOPAMINE RECEPTOR
INTERACTING PROTEIN homolog [Rattus norvegicus]
(SPTR|Q925G7, evidence: FASTY, 92.7%ID, 100%length,
match=2109)
putative"
/codon_start=1
/protein_id="BAC32320.1"
/db_xref="GI:26337271"
/translation="MAQKHPGERRLCAGHRSGGTSLTSGSVDPPEILFSLGLRDSAE
TAPNTRCLKHSGPKYTPPNPAHSDPSHGPPGPPRGPGVDPDESETGSESGV
DOELSENCTGYQEDGSPFELIPACNQCSPGVPEGTSEEGDSSSLCHRCTSP
ALGDEBLEBYDDEPKJFDFSVSSGKAPUSRRKRLIKEDVDSGRREPKA
PGRHLARSQTDKRGGLGWGVEELCOLQAGFWLIELLVGEVETCGHLIYA
CRKLKSGDLFRVWVGVARLGLGARMFQFQSQFFCVVGLLIRLRVVGAFLL
ALALFLGCLQWGRFVGLGNLRWDRDTAMLFSWLGSPALHCLTLKLDSPWQOLV
RLIOWGQELPWKQRTKKQGNAPVAGRYCQPEEVTLLTMAGVPDELPNPHVLG
VEATSDTELKAYRLQVAVHPDKNHPRAEEAFKILRAADIVSNPERKEVEMKR
MAENELSRVNEFLSKODDLEKAMTWMSRCQKRRFEMDEPKSARYCASCNRL
HPAEGDFWAESSLGLKLTITFALMDGKVYDITWAGCQVIGISPDTHRVPHVIFSGS
RVPGTSGRQATPPSPADLQDFLSIRIFQVPSGADVQWELLCRTSPWPDHFDL"

CDS

Db 1622 GTCCAGCAACCCAGAGAGCGGAGGAATATGAGATGAACCGGATGGCAGAGATGAGCTC 1681
QY 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsn 100
Db 1682 AGCCGTCAGTGAATGAGTTTCTGTCGAAATACAAAGATGACCTCAAGGAGCAATCAAC 1741
QY 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgPheGluMetAspArgGluPro 120
Db 1742 AGCATGATGTGCAGCAGATGCCAAGAAAGCATAGGAGGTTTGAAATGGACCGGAAACC 1801
QY 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140
Db 1802 AGAGTGCAGATACGTGCTGAGTGTAACAGCTGCATCTCTGCTGAGGAAGAGACTTT 1861
QY 141 TipAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160
Db 1862 TGGGCAGAGTTCGAGCATGCTGGGCCCTCAAGATCACTTACTTTCCGCTGATGATGGAAG 1921
QY 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyLysSerProAspThrHis 180
Db 1922 GTGTATGACATCACAGAGTGGCTGATGCCAGCGTGTGGGTATCTCCCCAGATCTCAC 1981
QY 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200
Db 1982 AGAGTTCTTACCACATCTCATTTGGTTCTCGGTACCCGCGACCACTGCGCGGACAGG 2041
QY 201 AlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220
Db 2042 GCCATCCAGAGTCCCTCTCTGCTGACCTGACGAGGATTTCTTGAGCGGATCTTTCAAGTA 2101
QY 221 -ProProGlyGlnMetProAsnGlyAsnPheAlaAlaProGlnProAlaProGlyAl 240
Db 2102 CCCTCCGGGGCGATGTCCAATGGAACTTTCTTTCGCGACCTCACTCCCTGGCCCTGGGAC 2161
QY 240 alalalalaserLysProAsnSerThrValProLysGlyGluAlaLysProLysArg 260
Db 2162 CACTTCGACCTCTAGCGCCCAACAGTTCAGTACCCCAAGGAGGAAGCAACCTAAACCGCG 2221
QY 260 gLysLysValArgArgProPheGlnArg 269
Db 2222 GAAGAAAGTGAGGGCGGCGCTTTTCAACGA 2249

RESULT 6

AK083878

LOCUS

DEFINITION

Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
enriched library, clone:D130038P18 product:DOPAMINE RECEPTOR
INTERACTING PROTEIN homolog [Rattus norvegicus], full insert
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

AK083878 4053 bp mRNA linear HTC 05-DEC-2002
Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
enriched library, clone:D130038P18 product:DOPAMINE RECEPTOR
INTERACTING PROTEIN homolog [Rattus norvegicus], full insert
sequence.

AK083878
AK083878.1 GI:26101555

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Alignment Scores:

Pred. No.: 2,25e-106 Length: 3022

Score: 1348.00 Matches: 251

Percent Similarity: 95.93% Conservative: 8

Best Local Similarity: 92.96% Mismatches: 10

Query Match: 94.07% Indels: 1

DB: 11 Gaps: 0

US-10-049-742-11 (1-269) x AK045345 (1-3022)

QY 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20

Db 1442 ATGCTGGGGTCTCTGAAGATGAATCAACCTTTTCATGCTGCGGGTTGAAGCTACA 1501

QY 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40

Db 1502 GCATCCGACACTGAACCTAAAGAGCGCTATAGGAGCTAGCAGTAATGGTCCATCCTGAT 1561

QY 41 LysAsnHisHisProArgAlaGluAlaGluAlaPheLysValLeuArgAlaAlaTipAspIle 60

Db 1562 AAAAAATCACCATCCCGGGCTGAGAGCGCTTCAAAATTTTGGCGGAGCTTGGACATT 1621

QY 61 ValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeu 80

Db 1621

TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohata,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
	20530913
	11076861
	4
TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staib,L.F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyohara,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,J., Kohsaki,S. and Hayashizaki,Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 695-690 (2001)
	21095660
	11217851
	5
TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
	6 (bases 1 to 4053)
	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,T., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akashira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp, URL:ftp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
	11076861
	4
TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp, URL:ftp://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:ftp://genome.gsc.riken.go.jp/ URL:ftp://fantom.gsc.riken.go.jp/
	11076861
	4
	1
TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Location/Qualifiers 1. .4053 /organism="Mus musculus" /mol_type="mRNA"
	11076861
	4
	1

misc_feature	94. .2205	/strain="C57BL/6J"	
		/db_xref="FANTOM DB:D130038P18"	
		/db_xref="taxon:10090"	
		/clone="D130038P18"	
misc_feature	94. .2205	/tissue type="spinal ganglion"	
		/clone_lib="RIKEN full-length enriched mouse cDNA library"	
		/dev_stage="12 days embryo"	
		/note="DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus] (SPTR Q925G7, evidence: FASTY, 92.7%ID, 100%length, match=2109)	
misc_feature	94. .2205	/putative"	
		BASE COUNT 1060 a 903 c 1039 g 1051 t	
		ORIGIN	
		Alignment Scores:	
misc_feature	94. .2205	Pred. No.: 2,73e-105 Length: 4053	
		Score: 1337.00 Matches: 251	
		Percent Similarity: 95.91% Conservative: 7	
		Best Local Similarity: 93.31% Mismatches: 11	
misc_feature	94. .2205	Query Match: 93.30% Indels: 1	
		DB: 11 Gaps: 0	
		US-10-049-742-11 (1-269) x AK083878 (1-4053)	
		Qy 1 MetAlaGlyValProGluAspGluLeuAsnProPheHisValLeuGlyValGluAlaThr 20	
misc_feature	94. .2205	Db 1397 ATGGCTGGGGTCTCTGAGATGAATGAACCTTTTCATGTCTGGGGTGTGAAGCTACA 1456	
		Qy 21 AlaSerAspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAsp 40	
		Db 1457 GCATCCGACACTGAACCTAAAGAGGCTATAGGCAGCTAGCAGTAATGGTCCATCTGAT 1516	
		Qy 41 LysAsnHisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTyrAspIle 60	
misc_feature	94. .2205	Db 1517 AAAATACCATCCCGGCTGAGGAGGCCCTTCAAAATTTTGGGGCAGCTTGGGACATT 1576	
		Qy 61 ValSerAsnAlaGluLysArgLysGluTyrGluValLysArgMetAlaGluAsnGluLeu 80	
		Db 1577 GTCCAGCAACCCAGAGAGCGGAAGGAATATGAGATGAACCGATGAGGAGGAGAGCTC 1636	
		Qy 81 SerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsn 100	
misc_feature	94. .2205	Db 1637 AGCCGGTCTAGTGAATGAGTTCTGTGCAAACTCAAGATGACCTCAAGAGGAGCAATGAAC 1696	
		Qy 101 ThrMetMetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluPro 120	
		Db 1697 ACATGATGTGACAGCAGATGCCAAGGAACCATAGGAGGTTTGAATGGACCGGAACCC 1756	
		Qy 121 LysSerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPhe 140	
misc_feature	94. .2205	Db 1757 AGAGTGGCCAGATACTGTGCTGAGTGTACAGGCTGCATCTCTGCTGAGGAGGAGACTTT 1816	
		Qy 141 TrpAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLys 160	
		Db 1817 TGGGAGAGTCGAGCATGCTGGGCCCTCAAGATCATTCTTTGGCTGTGGATGGAAG 1876	
		Qy 161 ValTyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHis 180	
misc_feature	94. .2205	Db 1877 GTGTATGACATCACAGAGTGGGCTGGATGCCAGCGTGTGGGTATCTCCCCAGATCTCAC 1936	
		Qy 181 ArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArg 200	
		Db 1937 AGATTCTCTTACCACATCTATTGTTCTTCGGGTACCCGACCCAGTGGCCGCGCAGAG 1996	
		Qy 201 AlaThrProAspAlaProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnVal 220	
misc_feature	94. .2205	Db 1997 GCCACTCCAGAGTC-CTCTCTGCTGACCTGCAGGATTTCTTGAGCCGGATCTTTCAAGTA 2055	
		Qy 221 ProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAla 240	
		Db 2056 CCTCCGGGGCGCATGTCCAATGGAACTCTTTTCCGGCCACTCACCTCTGGCCCTGGGACC 2115	

```

QY      241 AlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArg 260
      ::: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      2116 ACTTCGACCTCTAGGCCCAACAGTTTCAGTACCCAGGAGAGCCAAACCTTAAACGGCG 2175

QY      261 LysLysValArgArgProPheGlnArg 269
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      2176 AAGAAAGTCAGGCGGCCCTTTCAACGA 2202

RESULT 7
LOCUS   BU538214
DEFINITION   BU538214 942 bp mRNA linear EST 13-SEP-2002
IMAGE:6568738 5', mRNA sequence.
ACCESSION   BU538214
VERSION     BU538214.1 GI:22848655
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 942)
REFERENCE   NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2754 row: o column: 10
            High quality sequence stop: 680.
            Location/Qualifiers
              1..942
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6568738"
                /tissue_type="adrenocarcinoma, cell line"
                /lab_host="PH108 (phage-resistant)"
                /clone_lib="NIH_MGC_107"
                /note="Organ: Breast; Vector: pOTB7; Site 1: EcoRI;
                Site 2: XhoI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH MGC Library."
BASE COUNT   236 a 235 c 244 g 222 t
ORIGIN
Alignment Scores:
Pred. No.:      5,3e-101      Length:      942
Score:          1280.00      Matches:    239
Percent Similarity: 99.59%      Conservative: 1
Best Local Similarity: 99.11%      Mismatches:  1
Query Match:     89.32%      Indels:     0
DB:              13          Gaps:         0

US-10-049-742-11 (1-269) x BU538214 (1-942)

QY      29 AlaTyArgGlnLeuAlaValMetValHisProAspLysAsnHisProArgAlaGlu 48
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1 GCCTATAGACAGCTGGCAGTGATGTTTCATCTGACAAAAATCATCCCCGGGCTGAG 60

QY      49 GluAlaPheLysValLeuArgAlaLanProAspIleValSerAsnAlaGluLysArgLys 68
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      61 GAGGCGCTTCAAGGTTTTCGAGCAGCTTGGGACATGTGCAGCAATGCTGAAAGCGAAG 120

```

```

QY      69 GluTyArgGlnMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeu 88
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      121 GAGTATGAGATGAAACGAATGGCAGAGATGAGCTGAGCGGTCAATAGATTTCTG 180

QY      89 SerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGln 108
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      181 TCCAAAGCTGCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTGTAGCCGATGCCAA 240

QY      109 GlyLysHisArgArgPheGluMetAspArgGluProLysSerAlaArgTyCysAlaGlu 128
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      241 GGAACATAGGAGGTTTGAATGGACCGGACCTTAAGAGTCCAGATATCTGTGCTGAG 300

QY      129 CysAsnArgLeuHisProAlaGluGluArgPheThrAlaGluSerSerMetLeuGly 148
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      301 TGTATAGGCTGCATCTCTGAGGAGGAGACTTTTGGCAGAGTCAACGATGTTGGGC 360

QY      149 LeuLysIleThrTyrrPheAlaLeuMetAspGlyLysValTyrrAspIleThrGluTrpAla 168
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      361 CTCAGATCACCTACTTTGCACCTGATGGATGGAAAGGTGTATGACATCACAGAGTGGGCT 420

QY      169 GlyCysGlnArgValGlyIleSerProAspThrHisArgValProTyrrHisIleSerPhe 188
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      421 GGATGCCAGCGTGTAGGTATCTCCCCAGATATCCACAGAGTCCCTTATCATCTCATTT 480

QY      189 GlySerArgIleProGlyThrArgGlyArgGlnArgAlaThrProAspAlaProAla 208
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      481 GGTTCCTCGGATTCCAGGCACAGAGCGCGGAGAGCCACCCAGATGCCCTCTCTGCT 540

QY      209 AspLeuGlnAspPheLeuSerArgIlePheGlnValProGlyGlnMetProAsnGly 228
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      541 GATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCCGGAGATGCCCAATGGG 600

QY      229 AsnPheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerLysProAsnSer 248
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      601 AACTTCTTTGACGCTCTCAGCCTGCCCTGGAGCGCTGCAGCCTTAAGCCCCACAGC 660

QY      249 ThrValProLysGlyGluAlaLysProLysArgArgLysValArgArgProPheGln 268
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      661 ACAGTACCCAGGAGAGCAACCTAAGCGCGGAGAAAGTAGGAGGCCCTTCAA 720

QY      269 Arg 269
      ||||
Db      721 CGT 723

RESULT 8
LOCUS   BU538351
DEFINITION   BU538351 955 bp mRNA linear EST 13-SEP-2002
IMAGE:6568891 5', mRNA sequence.
ACCESSION   BU538351
VERSION     BU538351.1 GI:22848792
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 955)
REFERENCE   NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished
JOURNAL
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2755 row: e column: 19
            High quality sequence stop: 583.
            Location/Qualifiers
              1..955
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6568891"
                /tissue_type="adrenocarcinoma, cell line"
                /lab_host="PH108 (phage-resistant)"
                /clone_lib="NIH_MGC_107"
                /note="Organ: Breast; Vector: pOTB7; Site 1: EcoRI;
                Site 2: XhoI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH MGC Library."
BASE COUNT   236 a 235 c 244 g 222 t
ORIGIN
Alignment Scores:
Pred. No.:      5,3e-101      Length:      942
Score:          1280.00      Matches:    239
Percent Similarity: 99.59%      Conservative: 1
Best Local Similarity: 99.11%      Mismatches:  1
Query Match:     89.32%      Indels:     0
DB:              13          Gaps:         0

US-10-049-742-11 (1-269) x BU538214 (1-942)

QY      29 AlaTyArgGlnLeuAlaValMetValHisProAspLysAsnHisProArgAlaGlu 48
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1 GCCTATAGACAGCTGGCAGTGATGTTTCATCTGACAAAAATCATCCCCGGGCTGAG 60

QY      49 GluAlaPheLysValLeuArgAlaLanProAspIleValSerAsnAlaGluLysArgLys 68
      ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      61 GAGGCGCTTCAAGGTTTTCGAGCAGCTTGGGACATGTGCAGCAATGCTGAAAGCGAAG 120

```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:656891"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site:1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 239 a 241 c 261 g 213 t 1 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 1,57e-97 Length: 955
Score: 1240.00 Matches: 238
Percent Similarity: 98.76% Conservative: 1
Best Local Similarity: 98.35% Mismatches: 2
Query Match: 86.53% Indels: 2
DB: 13 Gaps: 0
```

US-10-049-742-11 (1-269) x B0538351 (1-955)

```
QY 29 AlaTyrArgGluLeuAlaValMetValHisProAspLysAsnHisProArgAlaGlu 48
DB 1 GCCTATAGACAGCTGCAGTGATGGTTTCCTCGACAAAAATCATCATCCCCGGGCTGAG 60
QY 49 GluAlaPheLysValLeuArgAlaAlaTyrAspIleValSerAsnAlaGluLysArgLys 68
DB 61 GAGGCCCTTCAAGGTTTTGGCAGCAGCTTGGGACATTGTGCAGCAATGCTGAAAGCGCAAG 120
QY 69 GluTyrGluMetLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeu 88
DB 121 GAGTATGAGATGAACGAATGGCAGAGAATGAGCTGAGCGGTCAGTAATAGATTCTTG 180
QY 89 SerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGln 108
DB 181 TCCAGCTGCAAGATGACCTCAAGGAGCAATGAATACTATGATGTGTAGCCGATCCCAA 240
QY 109 GlyLysHisArgArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGlu 128
DB 241 GGAAGCATAGGAGGTTTCAATGGACCGGAACTTAAGAGTSCCAGATACTGTGCTGAG 300
QY 129 CysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGly 148
DB 301 TGTATAGGCTGCATCTCTGCTGAGGAAGAGACTTTTGGGCAGAGTCAAGCATGTTGGGC 360
QY 149 LeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAla 168
DB 361 CTCAGATCACCCTACTTTGCATGTATGGATGGAAAGGTGTATGACATCACAGAGTGGGT 420
QY 169 GlyCysGlnArgValGlyIleSerProAspThrHisArgValProTyrHisIleSerPhe 188
DB 421 GGATGCCACCGTGTAGGTATCTCCACAGATATCCACAGAGTCCCTATCACATCTCATTT 480
QY 189 GlySerArgIleProGlyThrArgGlyArgGlnArgAlaThrProAspAlaProAla 208
DB 481 GGTTCCTCGGATTCACAGCACAGAGGCGGCAGAGAGCCACCCACAGATGCCCTCTCTGCT 540
QY 209 AspLeuGlnAspPheLeuSerArgIlePheGlnValProProGlyGlnMetProAsnGln 228
DB 541 GATCTTCAGGATTTCTTGATCGGATCTTTCAAGTACCCCGAGGGGAGATGCCCAATGG 600
QY 228 ValAsnPheAlaAlaProGlnProAlaProGlyAlaAlaAlaSerLysProAsnSe 248
DB 601 GAATCTTTTGAGCTCTTCAGCTCCCTCGAGCCGCTGAGCCGCTTAAGCCAAAC-AG 659
QY 248 rThrValProLysGlyGluAlaLysProLysArgArgLysLysValArgArgProPheGln 268
```

```
DB 660 CACAGTACCCAGGAGGAGGAGCAACCTAAGCGCGGAAAGAAAGTGTAGGAGGCCCTTCCA 719
QY 268 nArg 269
DB 720 CCGT 723
RESULT 9
LOCUS BM802135
DEFINITION AGENCOURT_6461654 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5559503
5', mRNA sequence.
ACCESSION BM802135
VERSION BM802135.1 GI:19118958
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 957)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LMNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12284 row: c column: 24
High quality sequence stop: 621.
FEATURES
Location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5559503"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 243 a 240 c 246 g 224 t 4 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 3,82e-96 Length: 957
Score: 1224.00 Matches: 230
Percent Similarity: 99.57% Conservative: 1
Best Local Similarity: 99.14% Mismatches: 1
Query Match: 85.42% Indels: 1
DB: 12 Gaps: 0
US-10-049-742-11 (1-269) x BM802135 (1-957)
```

```
QY 38 HisProAspLysAsnHisProArgAlaGluAlaPheLysValLeuArgAlaAla 57
DB 1 TATCTTGACAA-AATCATCATCCCCGGCTGAGGAGCCCTTCAAGGTTTTCGAGCAGCT 59
QY 58 TrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGlu 77
DB 60 TGGACATTGTGAGCAATGCTGAAAGCGAAGGAGTACGAGATGAAACGAATGGCAGAG 119
QY 78 AsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGlu 97
DB 120 AATGAGCTGAGCCGGTCAGTAAATGAGTTTCTGTCCAGAGCTGCAAGATGACCTCAAGGAG 179
```

Qy	98	AlaMetAsnThrMetMetCysSerSerArgCysGlnGlyLysHisArgArgPheGluMetAsp	117
Db	180	GCAATGAATACTATGATGTAGCCGATGCCAAGAAAGCATAGGAGGTTTGAAATGGAC	239
Qy	118	ArgGluProLysSerAlaAArgTyrCysalagLysCysAsnArgLeuHisProAlaGluGlu	137
Db	240	CGGGAACCTAAGATGCCAGATACTGTGCTGAGTGTAAATAGGCTGCATCCTCTCAGAGAA	299
Qy	138	GlyAspPheThrAlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMet	157
Db	300	GGAGACTTTTGGGCAGAGTCAAGCATGTTGGCCCTCAAGATCACCTACTTTGCATCTGATG	359
Qy	158	AspGlyLysValTyrAspIleThrdLutrpAlaglyCysGlnArgValGlyIleSerPro	177
Db	360	GATGGAAAGGTGTATGACATCACAGAGTGGGTGGATGCCAGCGTGTAGGTATCTCCCCA	419
Qy	178	AspThrHisArgValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGly	197
Db	420	GATACCACACAGAGTCCCTATCATCATCTCATTTGGTTCTCGATTCAGGCCACACAGAGG	479
Qy	198	ArgGlnArgAlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIle	217
Db	480	CGGCAGAGAGCCACCCAGATGCCCTCTGCTGATCTTCAGGATTTCTTGAGTCGGATC	539
Qy	218	PheGlnValProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAla	237
Db	540	TTTTCAGTACCCCCAGGGCAGATGCCCAATGGAACTTCTTTGCAGCTCTCTCAGCCTGCC	599
Qy	238	ProGlyAlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAlaLysPro	257
Db	600	CCTGGAGCCGCTGCAGCCCTCTAAGCCCCAACGACAGTACCCAAAGGGAGAGCCCAACCT	659
Qy	258	LysArgArgLysLysValArgArgProPheGlnArg	269
Db	660	AAGCGGGGAAGAAGTGCAGAGGCCCTTCCAACT	695

Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using Zap-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC library."

BASE COUNT	234 a	228 c	236 g	218 t
ORIGIN				
Alignment Scores:				
Pred. No.:	3 93e-93	Length:	918	
Score:	1189.00	Matches:	223	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	82.97%	Indels:	0	
DB:	13	Gaps:	0	
US-10-049-742-11 (1-269) x BU956189 (1-918)				
Qy	47	AlaGluGluAlaPheLysValLeuArgAlaAlaTtpAspileValSerAsnAlaGluLys	66	
Dd	3	GCTGAGGAGCCCTTCAGGTTTTGCGAGCAGCTTGCGACATTGTGCAGCAATGCTGAAAAG	62	
Qy	67	ArgLysGluTyrgLuMetLysArgMetAlaGluAenGluLeuSerArgSerValAenGlu	86	
Dd	63	CGAAAGGAGTATGAGATGAAGCGAATGGCAGAGAATGAGCTGAGCCGGTCAGTAATGAG	122	
Qy	87	PheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArg	106	
Dd	123	TTTTCTGTCAGGCTGCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTGATGCCGA	182	
Qy	107	CysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSerAlaArgTyrcys	126	
Dd	183	TGCCAAGGAAGCATAGGAGGTTTGAAATGGACCGGGAACCTTAAGAGTCCCAGACTCT	242	
Qy	127	AlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAlaGluSerSerMet	146	
Dd	243	GCTGAGTGTAATAGGCTGCATCTCTGTGAGGAAGGAGACTTTTGGGCAGAGTCAAGCATG	302	
Qy	147	LeuGlyLeuLysIleThrTyrrPheAlaLeuMetAspGlyLysValTyraSpileThrGlu	166	
Dd	303	TTGGGCGCTCAAGATCACCTACTTTTGACCTGTGATGGAAAGGTGTATGACATCACAGAG	362	
Qy	167	TtpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArgValProTyrrHisile	186	
Dd	363	TGGGCTGGATGCCAGCGTGTAGTATCTCCCGAGATACCACAGAGTCCCCCTATCACATC	422	
Qy	187	SerPheGlySerArgIleProglyThrArgGlyArgGlnArgAlaThrProaspAlapro	206	
Dd	423	TCATTTGGTTCTCGGATTCAGGCACCCAGAGGGCGGAGAGGCCACCCAGATGCCCT	482	
Qy	207	ProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnValProProGlydnMetPro	226	
Dd	483	CTGTGCTGATCTTCAGGATTCTTGAGTCGGATCTTTTCAAGTACCCCCAGGGCAGATGCC	542	
Qy	227	AsnGlyAsnPhePheAlaAlaproGlnProAlaProGlyAlaAlaAlaSerLysPro	246	
Dd	543	AATGGGAACCTCTTTTGACCTCTCTCAGCCTGCCCTTGAGGCCGCTGCAGCCCTTAAGCCC	602	
Qy	247	AsnSerThrValProLysGlyGluAlaLysProLysArgArgGlyLysValArgArgPro	266	
Dd	603	AACAGCACAGTACCACGAGGAGAACCCNAACCTTAGCGGCGGAAGAAAGTGAAGAGGCC	662	
Qy	267	PheGlnArg	269	
Dd	663	TTCCAACGT	671	
RESULT 11				
BG763500				
LOCUS	BG763500 878 bp mRNA linear EST 15-MAY-2001			
DEFINITION	602735591F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4860694 5', mRNA sequence.			

```

ACCESSION   BG763500
VERSION     BG763500.1  GI:14074153
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 878)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/BTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1718 row: f column: 23
            High quality sequence stop: 811.
FEATURES    Location/Qualifiers
            1..878
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4860694"
            /tissue_type="melanotic melanoma, high MDR (cell line)"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_49"
            /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGACAGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using ZAP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
            Library."
BASE COUNT  227 a 211 c 238 g 201 t
ORIGIN
Alignment Scores:
Pred. No.:      9,83e-89      Length:      878
Score:          1138.00      Matches:    213
Percent Similarity: 99.53%      Conservative: 0
Best Local Similarity: 99.53%      Mismatches: 1
Query Match:    79.41%      Indels:     0
DB:             12          Gaps:        0
US-10-049-742-11 (1-269) x BG763500 (1-878)
QY          56 AlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTrpGluMetLysArgMet 75
Db          2 GCAGCTGGGACATGTGCAGCAATCTGTAAGCGAAGAGGATGATGAGATGAACGATG 61
QY          76 AlaGluAsnGluSerArgSerValAsnGluPheLeuSerLysLeuGlnAspLeu 95
Db          62 GCAGAGAATGAGCTGAGCGCGTCAGTAATAGTTCTCTCCAAAGTCGCAAGATGACCTC 121
QY          96 LysGluAlaMetAsnThrMetMetCysSerArgCysGlnGlyHisArgGArgPheGlu 115
Db          122 AAGGAGGAATGAATACATATGATGTAGCCCATGCCAAGGAAGAAACATAGGAGGTTGAA 181
QY          116 MetAspArgGluProLysSerAlaArgTYrCysAlaGluCysAsnArgLeuHisProAla 135
Db          192 ATGGACCGGGAACCTTAGAGTCCAGACTACTGTCTGAGTGTAATAGGCTGATCTCTGCT 241
QY          136 GluGluGlyAspPheTrpAlaGluSerSerMetLeuGlyLeuLysIleThrTYrPheAla 155
Db          242 GAGGAAGGAGACTTTTGGGCAGAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTGCA 301

```

```

QY          156 LeuMetAspGlyLysValTrpAspIleThrGluTrpAlaGlyCysGlnArgValGlyIle 175
Db          302 CTGATGGATGGAAAGGTGTATGACATCAGAGTGGGCTGGATGCCAGCGGTGAGGTATC 361
QY          176 SerProAspThrHisArgValProTYrHisIleSerPheGlySerArgIleProGlyThr 195
Db          362 TCCCCAGATACCCACAGAGTCCCTTATCATCTCATTTGGTTCTCGGATTCAGGACACC 421
QY          196 ArgGlyArgGlnArgAlaThrProAspAlaProProAlaAspLeuGlnAspPheLeuSer 215
Db          422 AGAGGCGCGCAGAGAGCCACCCAGATGCCCTCTCTGCTGATCTTCAGGATTTCTTGAGT 481
QY          216 ArgIlePheGlnValProProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGln 235
Db          482 CGGATCTTTCAAGTACCCCGCAGGAGATGCCCAATGGGAACCTCTTTGCAGCTCCTCAG 541
QY          236 ProAlaProGlyValAlaAlaAlaSerLysProAsnSerThrValProLysGlyGluAla 255
Db          542 CTGCCCCCTGGAGCCGCTGACGCTCTTAAGCCCAACAGCAGACATACCCAGGAGAAAGCC 601
QY          256 LysProLysArgArgLysLysValArgArgProPheGlnArg 269
Db          602 AAANCTAAGCGCGGGAAGAAAGTGAGGAGGCCCTTCCACGT 643
RESULT 12
BQ691177
LOCUS      BQ691177
DEFINITION BQ691177 904 bp mRNA linear EST 15-JUL-2002
            AGENCOURT_8340957 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6249114
            5', mRNA sequence.
ACCESSION  BQ691177
VERSION     BQ691177.1  GI:21816493
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 904)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2390 row: a column: 19
            High quality sequence stop: 609.
FEATURES    Location/Qualifiers
            1..904
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6249114"
            /tissue_type="ductal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_110"
            /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGACAGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."
BASE COUNT  219 a 244 c 239 g 202 t
ORIGIN
Alignment Scores:
Pred. No.:      1.01e-88      Length:      904

```

Score:	1138.00	Matches:	215
Percent Similarity:	98.62%	Conservative:	0
Best Local Similarity:	98.62%	Mismatches:	2
Query Match:	79.41%	Indels:	1
DB:	13	Gaps:	0
US-10-049-742-11 (1-269) x BQ691177 (1-904)			
QY	53	ValLeuArgAlaIaIaTrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMet	72
Db	1	GTITTTGGAGCAGCTTGGGACATTGTGAGCAATCTGAGAAAGCGAAAGGAGTATGAGATG	60
QY	73	LysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeuGln	92
Db	61	AAACGAATGCAGAGAATGAGCTGAGCGGTGACGTAAATGAGTTCTGTCCAAAGCTCAA	120
QY	93	AspAspLeuLysGluAlaMetAsnThrMetCysSerArgCysGlnGlyLysHisArg	112
Db	121	GATGACCTCAAGGAGGCAATGATACATATGATGTGTAGCCGATGCTCCAAAGAAACATAGG	180
QY	113	ArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLeu	132
Db	181	AGGTTTGAATGGACCGGAACCTTAAGAGTGCCAGATCTGTGCTGAGTGTAAATAGGCTG	240
QY	133	HisProAlaGluGluGlyAspPheTrpAlaGluSerSerMetLeuGlyLeuLysIleThr	152
Db	241	CATCCTCTGCTGAGGAAGAGACTTTTGGCCAGAGTCAAGCATGTGTGGCCCTCAAGATCAC	300
QY	153	TyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTrpAlaGlyCysGlnArg	172
Db	301	TACTTTGCATGTGATGTAAGAGGTGTATGACATCACAGATGGGTGGATGCCAGCT	360
QY	173	ValGlyLeuSerProAspThrHisArgValProTyrHisIleSerPheGlySerArgIle	192
Db	361	GTAGGTATCTCCAGATACCCAGAGTCCCTTATCACATCTCATTTGGTCTCGGATT	420
QY	193	ProGlyThrArgGlyArgGlnArgAlaThrProAspAlaProAlaAspLeuGlnAsp	212
Db	421	CCAGGCACCAAGGGCGGAGAGAGCCACCCAGATGCCCTCTCTGCTGATCTTCAGAT	480
QY	213	PheLeuSerArgIlePheGlnValProProGlyValMetProAsnGlyAsnPhePheAla	232
Db	481	TTCTTGAGTCGATCTTTCAAGTACCCCGGGGAGATGCCCAATGGGAATCTTTTGCA	540
QY	233	AlaProGlnProAlaProGlyAlaAlaAlaAlaSerLysProAsnSerThrValProLys	252
Db	541	GCTCCTCAGCTGCCCTGGAGCGCTGAGCTCTTAAGCCCAACAGCACAGTACCCAG	600
QY	253	GlyGluAlaLysProLysArgArgLysLysValArgArg-ProPheGlnArg	269
Db	601	GGAGAAGTCAAACTAAGCGCGGAAACAAAGTGAGGAGGCCCTTCCAAAGT	652
RESULT 13			
BE779250			
LOCUS			
DEFINITION			
601464514F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3867794 5',			
mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLAM9614 row: 1 column: 03			
High quality sequence stop: 643.			
FEATURES			
source			
1. 723			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:3867794"			
/tissue_type="retinoblastoma"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH MGC 67"			
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: Not1;			
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.			
Average insert size 1.75 kb. Library constructed by Life			
Technologies."			
BASE COUNT			
ORIGIN			
201 a 162 c 208 g 152 t			
Alignment Scores:			
Pred. No.: 1.71e-83 Length: 723			
Score: 1076.50 Matches: 222			
Percent Similarity: 92.62% Conservative: 4			
Best Local Similarity: 90.98% Mismatches: 16			
Query Match: 75.12% Indels: 7			
DB: 10 Gaps: 2			
US-10-049-742-11 (1-269) x BE779250 (1-723)			
QY	23	AspValGluLeuLysLysAlaTyrArgGlnLeuAlaValMetValHisProAspLysAsn	42
Db	2	GATGTTGAACCTGAAGAGGCGCTTAGACAGCTGCGAGTGGTTCATCTCGACAAAT	61
QY	43	HisHisProArgAlaGluGluAlaPheLysValLeuArgAlaAlaTrpAspIleValSer	62
Db	62	CATCATCCCCGGCTGAGGAGGCGCTTCAAGGTTTTCGAGCAGCTTGGGACATTGTGAGC	121
QY	63	AsnAlaGluLysArgLysGluTyrGluMetLysArgMetAlaGluAsnGluLeuSerArg	82
Db	122	AATGCTGAAAGCGGAAAGGAGTATGAGATGAAACGAATGCGAGAGATGAGTCAGCGG	181
QY	83	SerValAsnGluPheLeuSerLysLeuGlnAspAspLeuLysGluAlaMetAsnThrMet	102
Db	182	TCAGTAAATGAGTTTCTGTCCAAAGCTGCAAGATGACCTCAAGGAGGCAATGAACTATG	241
QY	103	MetCysSerArgCysGlnGlyLysHisArgArgPheGluMetAspArgGluProLysSer	122
Db	242	ATGTGTAGCCGATGCCAAGAAAGCATAGGAGGTTTGAATGGACCGGAACTTAAGAT	301
QY	123	AlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrpAla	142
Db	302	GCCAGATACCTGCTCAGTGTAAATAGCTGCATCTCTGCTGAGGAGGAGACTTTTGGGCA	361
QY	143	GluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysValTyr	162
Db	362	GAGTCAAGCATGTTGGGCGCTCAAGATCACCTACTTTGCACTGATGGAAGGCTGAT	421
QY	163	AspIleThrGluTrpAlaGlyCysGlnArgValGlyLysSerProAspThrHisArgVal	182
Db	422	GACATCACAGAGTGGGCTGGATGCCAGCGTGTAGGTATCTCTCCAGATACCCAGAGTC	481
QY	183	ProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArgAlaThr	202
Db	482	CCCTATCATCTCATTTGGTTCTCGATTCCAGGCCAGAGGGCGGAGAGGCCACC	541
QY	203	ProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnValProPro	222
Db	542	CAGATTGCCCTCTCTGCTGATCTTCAGGATTTCTTGGTGGGATCTTTCAAGTACC-CCA	600
QY	223	GlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAlaAla	242

Db 601 GGGCAGATGCCAAT-GGGAACTCTTGGCAGTC---CTCAGCTGCCCTGGAGCCCGCTG-- 654

Qy 243 AlaSerLysProAsnSerThrValProLysGlyGluAlaLysProLysArgArgLysLys 262

Db 655 AGCTCTAAGCCAAAGGACAGTAA-CCCAAGGAGAAAG-CAACCT--AACGGGGAANA 709

Qy 263 ValArgArgPro 266

Db 710 GTGAGGAGGCC 721

RESULT 14

BM424147 927 bp mRNA linear EST 29-JAN-2002

LOCUS AGENCOURT 6398639 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5517411

DEFINITION 5', mRNA sequence.

ACCESSION BM424147

VERSION BM424147.1 GI:18392359

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 927)

TITLE NIH-MGC <http://mgs.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1CM2019 row: j column: 04

High quality sequence stop: 460.

FEATURES

source Location/Qualifiers

1..927

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5517411"

/tissue_type="amelanotic melanoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_41"

/note="Organ: skin; Vector: pOTS7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 236 a 241 c 245 g 204 t 1 Others

ORIGIN

Alignment Scores:

Pred. No.: 2,44e-83 Length: 927

Score: 1076.00 Matches: 214

Percent Similarity: 93.51% Conservative: 2

Best Local Similarity: 92.84% Mismatches: 9

Query Match: 72.09% Indels: 6

DB: 12 Gaps: 2

US-10-049-742-11 (1-269) x BM424147 (1-927)

Qy 33 LeuAlaValMetVal-HisProAspLysAsnHisProArgAlaGluGluAlaPheLy 52

Db 11 CTGCGAGTGATGGTTNCATCTGACAAAATCATCATCCCGGCTGAGGAGGCTTCAA 70

Qy 52 sValLeuArgAlaAlaTrpAspIleValSerAsnAlaGluLysArgLysGluTyrGluMe 72

Db 71 GGTTTTGGCAGCAGCTTGGGACATTGTGAGCAATGCTGAAAGCGAAAGAGATATGAGAT 130

Qy 72 tLysArgMetAlaGluAsnGluLeuSerArgSerValAsnGluPheLeuSerLysLeuG1 92

Db 131 GAACGAAATGCGACAGAAATGAGTGAGCGGTCACTAATGAGTTTCTGTCCAAGCTGCA 190

Qy 92 nAspAspLeuLysGluAlaMetAsnThrMetMetCysSerArgCysGlnGlyLysHisAr 112

Db 191 AGATGACCTCAAGGAGGCAATGAATACTATGATGTGTAGCCGATGCCAAGGAAAGCATAG 250

Qy 112 gArgPheGluMetAspArgGluProLysSerAlaArgTyrCysAlaGluCysAsnArgLe 132

Db 251 GAGGTTCGAAATGGACCGGAACTTAAGAGTGCCAGATACTGTGCTGAGTGTAAATAGGCT 310

Qy 132 uHisProAlaGluGluGlyAspPheTTPAlaGluSerSerMetLeuGlyLeuLysIleTh 152

Db 311 GCATCCTGCTGAGGAGGAGACTTTTGGGACAGATCAAGCATGTGGGCTCAAGATCAC 370

Qy 152 rTyrPheAlaLeuMetAspGlyLysValTyrAspIleThrGluTTPAlaGlyCysGlnAr 172

Db 371 CTACTTTGCACTGATGGATGAAAGGTGTATGACATCACAGATGGGCTGGATGCCAGCG 430

Qy 172 gValGlyLysSerProAspThrHisArgValProTyrHisLysSerPheGlySerArgI1 192

Db 431 TGTAGGTATCTCCCGAGATACCCACAGAGTCCCTTATCACATCTCATTTGGTTCTCGGAT 490

Qy 192 eProGlyThrArgGlyArgGlnArgAlaThrProAspAlaProProAlaAspLeuGlnAs 212

Db 491 TCCAGGACACAGAGCGCGGACAGAGCCACCCAGATGCCCTCTCTGATCTCAGGA 550

Qy 212 pPheLeuSerArgIlePheGlnValProProGlyGlnMet-ProAsnGlyAsnPhePheA 232

Db 551 TTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGGCCCATGGGAACTTCTTTG 610

Qy 232 laAla-ProGlnProAla-ProGlyAlaAlaAlaSer---LysProAsnSerThrVa 250

Db 611 CAGCTCCCTCAACCTGCCCCCTGGAGCGCGTGGAGCCCTCTTAAACCCCAACAGCCAGT 670

Qy 250 lProLysGlyGluAlaLysProLys 258

Db 671 ACCCCAGGGGAG---AAACCCAAA 692

RESULT 15

BF679562 874 bp mRNA linear EST 21-DEC-2000

LOCUS 602153994F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4295182 5',

DEFINITION mRNA sequence.

ACCESSION BF679562

VERSION BF679562.1 GI:11953457

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 874)

TITLE NIH-MGC <http://mgs.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: CLONETECH Laboratories, Inc.

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1CM1145 row: k column: 23

High quality sequence stop: 697.

FEATURES

source Location/Qualifiers

1..874

/organism="Homo sapiens"

US-10-049-742-11 (1-269) x BP679562 (1-874)

Qy	62	SerAsnAlaGluLysArgGlyGluTyrGluMetLysArgMetAlaGluAsnGluLeuSer	81
Db	2	AGCAATGCTGAAAGCGAAAGAGTATGAGATGAAACGAATGCAGAGAATCAGCTGAGC	61
Qy	82	ArgSerValAsnGluPheLeuSerLysLeuGlnAspLeuLysGluAlaMetAsnThr	101
Db	62	CGGTTCAGTAAATGAGTTCTGTGCCAAGTGCAGAGTACCTCAAGGAGGCAATGAATACT	121
Qy	102	MetMetCysSerArgCysGlnGlnLysHisArgArgPheGluMetAspArgGluProLys	121
Db	122	ATGATGTGTAGCCGATGCCAGGAAGCATAGGAGGTITGAATGCACCGGAACCTTAG	181
Qy	122	SerAlaArgTyrCysAlaGluCysAsnArgLeuHisProAlaGluGluGlyAspPheTrp	141
Db	182	AGTCCACGATACTGTGCTGAGTGTAAATAAGCTGCATCCTGCTGAGGAAGGAGACTT-TGG	240
Qy	142	AlaGluSerSerMetLeuGlyLeuLysIleThrTyrPheAlaLeuMetAspGlyLysVal	161
Db	241	GCAGAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTGCACGTGATGGATGGAAAGGGT	300
Qy	162	TyrAspIleThrGluTrpAlaGlyCysGlnArgValGlyIleSerProAspThrHisArg	181
Db	301	TATGACATCACAGAGTGGCTGGATGCCAGCGTGTAGGTATCTCCCCAGATACCCACAGA	360
Qy	182	ValProTyrHisIleSerPheGlySerArgIleProGlyThrArgGlyArgGlnArgAla	201
Db	361	GTCCCCATCATCATCTCATTTGGTTCTCGGATCCAGGACCAGAGGGCGGAGAGAGCC	420
Qy	202	ThrProAspAlaProProAlaAspLeuGlnAspPheLeuSerArgIlePheGlnValPro	221
Db	421	ACCCACAGATGCCCTCCCTGCTGATCTTCAGGATTTCTTGAGTCGATCTTTCAAGTAGCCC	480
Qy	222	ProGlyGlnMetProAsnGlyAsnPhePheAlaAlaProGlnProAlaProGlyAlaAla	241
Db	481	CCAGGGCAGATGCCCAATGGGAACCTCTTTGCAGCTCTCTCAGCCTGCCCTCGAGCCACT	540
Qy	242	AlaAlaSerLysProAsnSerThrVal-ProLysGlyGluAlaLysProLysArgArgly	261
Db	541	GCAGCTCTAACCCCAACAGACACACAGTTCTCCTACGGGAGAGGCAAACTCTACGGCGGAA	600
Qy	261	LysValArgArgProphe	267
Db	601	GAAAGTGAGGAGGCCCTTC	619